





UNIVERSITY OF ILLINOIS AT  
CHICAGO  
801 SO. MORGAN  
CHICAGO, IL 60607



This book is the gift of

Dreyfus Fund

UNIVERSITY of ILLINOIS







Q

1

A773

V. 2

1949

PER



AUSTRALIAN JOURNAL  
OF  
SCIENTIFIC RESEARCH

SERIES A  
PHYSICAL SCIENCES

VOLUME 2

MELBOURNE  
1949



# AUSTRALIAN JOURNAL OF SCIENTIFIC RESEARCH

Published by the Commonwealth Scientific and Industrial Research Organization and the Australian National Research Council

Issued in two series :

**Series A—Physical Sciences**

(Four issues each year, appearing March, June, September, December)

**Series B—Biological Sciences**

(Four issues each year, appearing February, May, August, November)

Price : Each series 30/- per annum, separate issues 7/6 each

---

## EDITORIAL BOARD

Chairman and Editor : N. S. Noble

Members : Professor F. M. Burnet

Professor W. J. Dakin

Professor E. J. Hartung

Professor L. H. Martin

Professor J. G. Wood

---

All enquiries and manuscripts should be forwarded to :

The Editor,

Australian Journal of Scientific Research,

Commonwealth Scientific and Industrial Research Organization,

314 Albert Street, East Melbourne, C.2, Victoria.



## CONTENTS

NUMBER 1, MARCH 1949

	PAGE
On the Practical Determination of Lunar and Luni-Solar Daily Variations in Certain Geophysical Data. By K. K. Tschu ..	1
Angular Distribution of $\alpha$ -Particles from the $\text{Li}^7(p, \alpha)\text{He}^4$ Reaction. By L. H. Martin, J. C. Bower, D. N. F. Dunbar, and F. Hirst ..	25
Hall Currents and the Ejection of Prominences by Sunspots. By R. G. Giovanelli .. .. .	39
Solar Radiation at 1200 Mc/s., 600 Mc/s., and 200 Mc/s. By F. J. Leahy and D. E. Yabsley .. .. .	48
Microwave Thermal Radiation from the Moon. By J. H. Piddington and H. C. Minnett .. .. .	63
Experiments in Seeding Cumuliform Cloud Layers with Dry Ice. By E. J. Smith .. .. .	78
The Relative Stability of Internal Metal Complexes. I. Complexes of 8-Hydroxyquinoline, Salicylaldehyde, and Acetylacetone. By L. E. Maley and D. P. Mellor .. .. .	92
The Iodination of Aromatic Compounds. I. The Action of Iodine on Aniline in the Presence of Mercuric Oxide. By L. Jurd .. .. .	111
Studies on the Lignin of <i>Eucalyptus regnans</i> . III. The Isolation of an Alkali Lignin from Sulphate Pulp. By J. W. T. Merewether ..	117
The Chemical Constituents of Australian <i>Flindersia</i> Species. I. Collinin-7-Geranoxy-8-Methoxycoumarin. By F. A. L. Anet, F. R. Blanks, and G. K. Hughes .. .. .	127
The Occurrence of $\beta$ -Carotene and "Phytofluene" in the Wood of <i>Acacia acuminata</i> . By P. Stanley and V. M. Trikojus .. ..	132

## NUMBER 2, JUNE 1949

	PAGE
The Position and Probable Identification of the Source of Galactic Radio-Frequency Radiation Taurus-A. By J. G. Bolton and G. J. Stanley	139
Experimental Designs Balanced for the Estimation of Residual Effects of Treatments. By E. J. Williams	149
The Wave Equations for Electromagnetic Radiation in an Ionized Medium in a Magnetic Field. By K. C. Westfold	169
The Structure of Cosmic Ray Air Showers. By J. R. Prescott and C. B. O. Mohr	184
Solar Radio-Frequency Radiation of Thermal Origin. By J. L. Pawsey and D. E. Yabsley	198
Bursts of Solar Radiation at Metre Wavelengths. By Ruby Payne-Scott	214
The Noise-like Character of Solar Radiation at Metre Wavelengths. By Ruby Payne-Scott	228
The Artificial Stimulation of Precipitation by Means of Dry Ice. By P. Squires and E. J. Smith	232
The Iodination of Aromatic Compounds. II. The Iodination of Aromatic Ethers and of 2-Naphthol. By L. Jurd	246
Alkaloids of the Australian Rutaceae: <i>Melicope fareana</i> . I. Isolation of the Constituent Alkaloids. By J. R. Price	249
Alkaloids of the Australian Rutaceae: <i>Melicope fareana</i> . II. Preliminary Examination of Melicopine, Melicopidine, and Melicopicine. By W. D. Crow and J. R. Price	255
Alkaloids of the Australian Rutaceae: <i>Melicope fareana</i> . III. Oxidative Demethylation of Melicopicine. By W. D. Crow	264
Alkaloids of the Australian Rutaceae: <i>Melicope fareana</i> . IV. Some Reactions of 1-Methyl-4-Quinolone-3-Carboxylic Acid, a Degradation Product of the Alkaloids. By J. R. Price	272
Alkaloids of the Australian Rutaceae: <i>Melicope fareana</i> . V. The Structure of the Alkaloids. By W. D. Crow and J. R. Price	282



## NUMBER 3, SEPTEMBER 1949

	PAGE
Graphical Study of the Dispersion of Electro-Magneto-Ionic Waves. By V. A. Bailey and J. A. Roberts .. .. .	307
Transients in an Ionized Medium with Applications to Bursts of Solar Noise. By J. C. Jaeger and K. C. Westfold .. .. .	322
The Distribution of Light from Optical Systems. By W. H. Steel ..	335
The Uniform Torsion of an Incomplete Tore. By W. Freiburger ..	354
Theoretical Aspects of Cloud Drop Distributions. By E. B. Kraus and Betty Smith .. .. .	376
The Contraction of Nuclear Emulsions. By S. L. Martin .. ..	389
The Second Viscosity Coefficient in Rheological Systems. By R. C. L. Bosworth .. .. .	394
The Dielectric Properties of Binary Systems of Ketones and Hydrocarbons. By R. J. Meakins .. .. .	405
Alkaloids of the Australian Rutaceae: <i>Acronychia Baueri</i> . I. The Isolation of the Alkaloids. By F. N. Lahey and W. C. Thomas ..	423
The Isolation of <i>d</i> -Sparteine from <i>Hovea</i> Species. By J. H. Morrison and K. G. Neill .. .. .	427
Alkaloids of the Australian Rutaceae: <i>Evodia xanthoxyloides</i> . I. Evoxanthine. By G. K. Hughes and K. G. Neill .. ..	429

## NUMBER 4, DECEMBER 1949

	PAGE
Diffusion in a Field of Homogeneous Turbulence. I. Eulerian Analysis. By G. K. Batchelor . . . . .	437
The Fully Developed Wake of a Circular Cylinder. By A. A. Townsend. .	451
The Uniform Flexure of an Incomplete Tore. By W. Freiburger and R. C. T. Smith . . . . .	469
On the Solution of the Equilibrium Equations of Elasticity in General Curvilinear Coordinates. By W. Freiburger . . . . .	483
Cosmic Ray Measurements between Australia and Japan. By P. G. Law, C. D. McKenzie, and H. D. Rathgeber . . . . .	493
Measurements of Solar Radiation at a Wavelength of 50 Centimetres during the Eclipse of November 1, 1948. By W. N. Christiansen, D. E. Yabsley, and B. Y. Mills . . . . .	506
Solar Radiation at a Wavelength of 10 Centimetres including Eclipse Observations. By J. H. Piddington and J. V. Hindman . . . . .	524
Solar Radiation of Wavelength 1.25 Centimetres. By J. H. Piddington and H. C. Minnett . . . . .	539
A Remarkable Thunderstorm Flight Record. By U. Radok . . . . .	550
Bond Localization Energy. I. Definition, Methods of Computation, and Relation to Bond Order. By R. D. Brown . . . . .	564
The Relative Stability of Internal Metal Complexes. II. Metal Derivatives of 8-Hydroxyquinoline 5-Sulphonic Acid and a Series of Monocarboxylic Mono- $\alpha$ -Amino Acids including Histidine. By L. E. Maley and D. P. Mellor . . . . .	579
The Iodination of Aromatic Compounds. III. Halogenation of Aromatic Ethers in the Presence of Hydrogen Peroxide. By L. Jurd. . . . .	595
Studies on the Lignin of <i>Eucalyptus regnans</i> . IV. Alkali Lignin. By J. W. T. Merewether . . . . .	600
The Chemical Constituents of Australian <i>Flindersia</i> Species. II. Braylin and Brayleyanin By F. A. L. Anet, G. K. Hughes, and E. Ritchie . . . . .	608
A Synthesis of Hygrine and of Cuscohygrine. By E. F. L. J. Anet, G. K. Hughes, and E. Ritchie . . . . .	616
Alkaloids of the Australian Rutaceae: <i>Acronychia Baueri</i> . II. Some Reactions of the Alkaloid Acronycine. By R. D. Brown, L. J. Drummond, F. N. Lahey, and W. C. Thomas . . . . .	622
Alkaloids of the Australian Rutaceae: <i>Acronychia Baueri</i> . III. The Structure of Acronycine. By L. J. Drummond and F. N. Lahey . . . . .	630
Index to Volume 2 . . . . .	639









# ON THE PRACTICAL DETERMINATION OF LUNAR AND LUNI-SOLAR DAILY VARIATIONS IN CERTAIN GEOPHYSICAL DATA

By K. K. TSCHU\*

Manuscript received October 27, 1948]

## Summary

This is a continuation of two earlier papers, in which Chapman and Miller (9, cf. Miller 13) developed the mathematical theory of a convenient method for the determination of the lunar daily variations in geomagnetic and meteorological elements, and for estimating the probable error of any determination thus made. Important practical details of the work are given in the present paper, with examples of the calculations on suitable computing forms. It also includes a new application of the Chapman-Miller Method to the determination of a luni-solar periodic component first predicted and investigated in ionospheric data by Martyn(7).

## 1. INTRODUCTION

The influence of the moon on the oceans has long been studied in detail ; the lunar tide is the outstanding regular oceanic variation, along with its smaller companion, the solar tide.

The moon produces a lunar tide in the atmosphere also, but there it is a minor phenomenon ; irregular changes due to atmospheric disturbances, and the thermal and tidal influence of the sun (in part magnified by resonance) much exceed the lunar tide. Nevertheless this tide is of great interest because of the simplicity and regular distribution of the tidal force. The tide itself may not be altogether simple because the atmosphere on which the forces act is not uniform over the earth ; but in unravelling the complex changes of the atmosphere it is of value to examine what this specially simple phenomenon has to teach us, even though the tide is so small as to make its detection rather difficult and laborious.

The tide manifests itself directly in a lunar semi-diurnal variation of barometric pressure(1), also of air temperature(2), and wind-components(3) ; as yet the latter variations have been little studied.

Appleton and Weekes(4) have shown that the effect of the lunar tide in the ionosphere is large, as shown by the lunar semi-diurnal variation of the height of the *E* layer.

The rise and fall of the air modifies the level at which the bulk of cosmic ray mesons are formed, and as these are unstable the number received at the ground increases or decreases according as their level of formation falls or rises ; thus there is a lunar semi-diurnal variation in the number of mesons observed at ground level, as Duperier(5) has shown.

\* The Queen's College, Oxford.

Besides these purely lunar variations, the lunar tide produces other phenomena in which the sun also plays a part. Chief among these luni-solar changes are the ionospheric electric currents induced by the lunar tide in the ionosphere owing to the tidal motion of the air across the earth's magnetic field; these reveal themselves by variations of the geomagnetic field (in all three elements) observable at the ground, and also in the earth potentials and earth currents induced thereby(6). These variations have a true lunar semi-diurnal component together with other components whose arguments involve both lunar and solar time, with periodicities approximating to one day, a third of a day, and so on.

As Martyn(7) has shown, the ionospheric electric currents are associated also with luni-solar periodic variations of height and electron density in the ionospheric layers, due to the force exerted by the geomagnetic field on the electric currents.

Lunar tidal movements are observable also in the solid body of the earth(8). Doubtless other geophysical elements are subject to detectable lunar or luni-solar daily variations which remain to be discovered.

## 2. METHODS OF COMPUTATION

This paper is concerned with the determination of lunar and luni-solar daily variations in geophysical elements in cases where these variations are overlaid by larger solar daily and irregular variations. This applies to all those phenomena mentioned in Section 1 whose main cause is the lunar atmospheric tide. Numerous methods of determination have been used with success in the past. As a result of long experience Chapman and Miller(9) have developed a method which they consider to be generally most suitable for routine application where an adequate amount of solar hourly or bi-hourly data is available, although other methods may be necessary or advantageous in special cases(10, 11, 12). The theory and nature of the Chapman-Miller method have been fully described by its authors(9); the present paper supplies a straightforward account of the practical applications of the method, with examples of the calculations on suitable computing forms.

In the course of the work the harmonic components of the solar daily variation (Section 9) are computed for all the groups of days for which the lunar or luni-solar harmonics are determined; it is advantageous to have these two sets of results for intercomparison.

The practical details here given include also the determination of the probable errors of the results, by methods developed by Miller(13); this is an important part of the work and should never be omitted.

In Section 12 some considerations are described which may usefully be taken into account when planning an investigation of the kind here described.

## 3. THE MATERIAL

The material used in the Chapman-Miller method consists of solar hourly, or bi-hourly, or tri-hourly data for the geophysical element considered. The individual values may be instantaneous values, or means for successive intervals



of one solar hour, two hours, or three hours ; these means are naturally taken to refer to the *central epoch* of the interval. The epochs to which the values refer are thus *separated* by intervals of exactly one, two, or three mean solar hours ; they may be exact hours of Greenwich, or standard, or true local time, or may differ from such times by any constant interval.

It is generally advantageous to use bi-hourly rather than hourly data : this practice, as compared with the use of hourly data, nearly halves the major routine labour involved, with only a slight loss of accuracy in the results ; this is because the alternate hourly values which are omitted when bi-hourly values are used are highly correlated with their neighbouring hourly values.

The subtraction of a *constant* from all the values used clearly will *not affect* any periodic variation to be determined. It is generally possible and convenient to represent each value by not more than three digits, after subtraction of a suitable constant from all values. The constant need not be the same throughout the whole of an investigation, but if it is changed at any stage in the series of data, care must be taken that *no discontinuity* results in any daily sequence used (Section 4).

Sometimes the values available are given to one more decimal place than necessary for the present purpose ; if so, it is generally sufficient to ignore the last figure completely without troubling to raise the preceding digit by one unit when the omitted digit is five or more.

Care must be taken in regard to *incomplete* daily sequences. When the gap is brief (e.g. only one value, or two) it may be possible to fill it in by interpolation. If this is not possible the whole of the corresponding daily sequence (Section 4) should be omitted.

In some investigations it is desirable to omit certain daily sequences, because they include so much irregular disturbance as to be likely to vitiate the determination of the small periodic variations sought. This is a matter for consideration in planning the work.

#### 4. THE DAILY SEQUENCES

In general the data used will be *given* in monthly tables, with those for each day occupying one row (or column), e.g. for solar hours  $H'$ ,  $H'+1$ , . . . .  $H'+23$  if hourly values are tabulated, or  $H'$ ,  $H'+2$ , . . . .  $H'+22$  if bi-hourly, or  $H'$ ,  $H'+3$ , . . . .  $H'+21$  if tri-hourly ;  $H'$  may have any value, but it is generally 0 or 1. In any case the values *used* for any day will form such a series, numbering  $S$  terms, where  $S$  will be 24, 12, or 8 according as hourly, bi-hourly, or tri-hourly values are used. An important feature in the Chapman-Miller method is that in the computations each *daily sequence* used includes  $S+1$  values ; the last of these is the value for the initial hour  $H'$  of the succeeding day, that is, it follows the  $S$ -th value for the current day at the interval  $24/S$  hours. The sequence will therefore consist of a set of numbers which may be denoted by  $f_s$  ( $s=0$  to  $s=S$ ), i.e.  $f_0, f_1, \dots f_S$ .

Each such daily sequence is to be regarded as a *unit*, which can be combined with all the other daily sequences, or with only some selected group of them. It is *characterized* by its date, namely, its year  $\alpha$ , its calendar month  $\beta$ , and its

number in the month  $\gamma$ . Alongside it (e.g. at the beginning of the sequence) must be placed the number  $\mu$  (associated with the lunar age or phase) corresponding to that date. These numbers ( $\mu=0$  to  $\mu=23$ ) were first introduced by Schmidt(14) in 1928 and were tabulated by Bartels and Faiselau(15) in their *Geophysikalischer Mond-Tafeln* for every Greenwich noon from 1850 to 1975. They are a *decreasing* measure of the moon's age to the nearest whole hour;  $\mu$  is zero or 24 at new moon, and decreases from 24 to 0 in the interval from one new moon to the next.

Each daily sequence may also be assigned other "character" figures  $C, C', C'', \dots$ , to be used in the selection of sets of sequences for separate treatment. These figures may, for example, be assigned on the basis of one or more of such "quantities" as the lunar distance, or the daily range of the geophysical element on the day in question, or the daily sunspot number, or the magnetic character figure and so on (Section 12 (b)).

Thus a *complete* (characterized) daily sequence will consist of the numbers  $\alpha, \beta, \gamma, \mu, C, C', C'', \dots$ , followed by  $S+1$  numbers  $f_s$  ( $s=0$  to  $s=S$ ) of the element for the day;  $C, C', C'', \dots$  are the character figures for the different daily "quantities" considered.

These complete daily sequences may be formed by extending the rows (or columns) of the original tables of the data by adding to each row (or column) the preliminary numbers  $\alpha, \beta, \gamma, \mu, C, C', C'', \dots$ , and the final number, the  $(S+1)$ -th value, for hour  $H'$  of the next day. If more convenient, the complete daily sequences may be written on separate sheets, by copying from the original tables.

If punched-card (e.g. Hollerith) equipments are available each complete daily sequence may be entered on a card or cards, giving complete freedom in the method of grouping daily sequences. For example, a 45-column card (12 rows in each column) could be used to give 13 bi-hourly values each of 3 digits, leaving 6 columns free for  $\alpha, \beta, \gamma$  (one digit each),  $\mu$  (two digits), and one single-digit character number  $C$ . Cards with more columns (e.g. the Hollerith 80-column cards), or more than one card per day, are needed if several "character" selections of the data are desired. When the data are entered on cards it is desirable that all the entries should be numbers *with the same sign*.

It should be *emphasized* that the use of punched-card equipments is only a (great) convenience, but not a necessity, for the application of the Chapman-Miller method.

When punched cards are not available, sets of selected daily sequences can be separated out from the tables of sequences by cutting up the tables into strips, one for each sequence, and then laying these together so that corresponding numbers in different sequences are brought alongside each other\*; this arrangement is, however, in general not very convenient to handle. A better plan is to keep the tables intact, and either to lay obscuring strips over the daily sequences not to be included in the set,\* or to fold the sheets so that

\* A sheet of glass can be laid over all to keep the strips or obscuring strips in place.

the unwanted sequences are hidden in the folds, only the wanted sequences remaining visible.

The subsequent analysis of the daily sequences is the same, whether the whole of the sequences are included in the investigation, or only a restricted set of them. Such a set may, for instance, include all those for a given decade, or all those for a given calendar month or season\* in a group of years, or all those that correspond to any particular daily character to which figures have been assigned. In any such case the selection of the sequences in the set will correspond to the choice of those for which a particular one of the preliminary figures  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $C$ ,  $C'$ ,  $C''$  . . . . (other than  $\mu$ ) has a chosen value; or the selection may require more than one of these characters to have given values, e.g. all those January days for which the sunspot character number had a given value. If punched cards are used, the selection of such sets is done very rapidly and conveniently by the sorting machine, sorting on the appropriate column or columns.

## 5. THE TWELVE LUNAR-AGE GROUPS AND THEIR SUMMATION

Any such collection or *set* of daily sequences is first divided into twelve groups according to the value of the number  $\mu$  for each sequence. As  $\mu$  is a (decreasing) measure of the lunar age, these are called *lunar-age groups*. The groups are numbered 0 to 11; the sequences assigned to the group with number  $r$  are those for which  $\mu=r$  together with those for which  $\mu=r+12$ . Let  $N_r$  denote the number of the sequences (in the set considered) in the group  $r$ . Then

$$\sum_{r=0}^{r=11} N_r = N, \dots\dots\dots (1)$$

where  $N$  denotes the total number of sequences in the collection or set (see Section 11 (e)).

For each group of sequences the values  $f_s$  for the same  $s$  are summed. If the sequences are entered on punched cards, the cards are summed for each of the series of 3 columns representing the  $s$ -th 3-digit number  $f_s$ . If the daily sequences are written in rows on sheets, the particular group of sequences considered is isolated in one of the ways suggested in Section 4, or otherwise, and sums of columns for each hour are formed. In this way we obtain a *group-sum sequence* of  $S+1$  sums, which for group  $r$  may be denoted by  $g_{s,r}$  ( $s=0$  to  $s=S$ ), derived from  $N_r$  days.

The addition of these group-sums for each value of  $s$  gives a *total-sum sequence* of sums  $g_{s,N}$

$$g_{s,N} = \sum_{r=0}^{r=11} g_{s,r} \quad (s=0 \text{ to } s=S), \dots\dots\dots (2)$$

which is the sequence of sums for the whole of the  $N$  sequences in the set considered (i.e.  $\sum_1^N f_s$  for each value of  $s$ ). The sums  $g_{s,r}$  and  $g_{s,N}$  should be

\* The seasonal sets are usually defined as follows:

*J* set = May, June, July, August (near the June solstice);

*E* set = March, April, September, October (near the equinoxes);

*D* set = November, December, January, February (near the December solstice).



independently calculated, and checked by verifying the last equation for each value of  $s$  (Section 11 (e)).

Clearly, these twelve group-sum sequences and the total-sum sequence of the set considered, each with  $S+1$  sums, are either hourly ( $S=24$ ), bi-hourly ( $S=12$ ), or tri-hourly ( $S=8$ ), in accordance with the individual daily sequences involved (Section 4).

## 6. THE PRIMARY HARMONIC ANALYSES

The group-sum sequence  $g_{s,r}$  ( $s=0$  to  $s=S$ ) for each group  $r$  is next harmonically analysed according to the following formulae

$$A_{p,r} = \sum_{s=0}^S g_{s,r} \cos (2\pi s p/S) + \frac{1}{2}(g_{S,r} - g_{0,r}), \quad \dots \dots \dots (3)$$

$$B_{p,r} = \sum_{s=0}^S g_{s,r} \sin (2\pi s p/S) + \frac{1}{2}(g_{S,r} - g_{0,r}) \cot (\pi p/S), \quad \dots \dots \dots (4)$$

which allow for the average non-cyclic change of the element from the beginning ( $s=0$ ) to the end ( $s=S$ ) of each day.

6. (a) The numbers  $A_{p,r}$ ,  $B_{p,r}$  are calculated *only* for  $p=2$  in cases where only a pure lunar semi-diurnal variation is expected or sought in the data, e.g. in air pressure, air temperature, wind-components, or cosmic ray data. This gives 24 numbers, namely, 12 of  $A_{2,r}$  and 12 of  $B_{2,r}$ .

6. (b) In analyses of geomagnetic, or earth-current, or earth-potential data, where luni-solar variations may be expected as well as the purely lunar tidal semi-diurnal variation, the analyses are usually made for  $p=1, 2, 3, 4$ . This gives 96 numbers  $A_{p,r}$  and  $B_{p,r}$ .

6. (c) In addition the same analyses are applied to the total-sum sequences,  $g_{s,N}$ , giving corresponding numbers  $A_{p,N}$ ,  $B_{p,N}$ , namely, 2 in the case of 6 (a) and 8 in the case of 6 (b). These independent analyses are checked according to the equations.

$$A_{p,N} = \sum_{r=0}^{11} A_{p,r}, \quad B_{p,N} = \sum_{r=0}^{11} B_{p,r}, \quad \dots \dots \dots (5)$$

for each value of  $p$  (see Sections 9 and 11 (e)).

## 7. THE SECONDARY HARMONIC ANALYSES

The 12-term sequences  $N_r$ ,  $A_{p,r}$ ,  $B_{p,r}$  for  $r=0$  to  $r=11$  are next likewise subjected to simple harmonic analyses to determine their main harmonic component according to the formulae

$$N_{1,A} = \sum_{r=0}^{11} N_r \cos (2\pi r/R), \quad N_{1,B} = \sum_{r=0}^{11} N_r \sin (2\pi r/R) \dots \dots \dots (6)$$

$$A_{p,A} = \sum_{r=0}^{11} A_{p,r} \cos (2\pi r/R), \quad A_{p,B} = \sum_{r=0}^{11} A_{p,r} \sin (2\pi r/R) \dots \dots \dots (7)$$

$$B_{p,A} = \sum_{r=0}^{11} B_{p,r} \cos (2\pi r/R), \quad B_{p,B} = \sum_{r=0}^{11} B_{p,r} \sin (2\pi r/R) \dots \dots \dots (8)$$

where  $R=12$ , the number of lunar-age groups. Note also that these three sets of formulae have the same form.

In the case of Section 6 (a) this gives 6 numbers only,  $N_{1, A}$ ,  $N_{1, B}$ ,  $A_{2, A}$ ,  $A_{2, B}$ ,  $B_{2, A}$ ,  $B_{2, B}$ . In the case of Section 6 (b) there are four of each of the numbers  $A_{p, A}$ ,  $A_{p, B}$ ,  $B_{p, A}$ ,  $B_{p, B}$  for  $p=1, 2, 3, 4$ , together with two numbers,  $N_{1, A}$  and  $N_{1, B}$ .

### 8a. THE FINAL REDUCTIONS; LUNAR SEMI-DIURNAL TERM ONLY

In such cases as are mentioned in Section 6 (a), where only the lunar semi-diurnal variation is expected and sought, the following combinations of the results described in Section 7 are formed; the formulae are given for any value of  $p$ , but we are here concerned *only with the value*  $p=2$ .

$$U_p = (A_{p, A} - B_{p, B}) - (A_{p, N} \cdot N_{1, A} - B_{p, N} \cdot N_{1, B})/N. \quad \dots\dots (9)$$

$$V_p = (B_{p, A} + A_{p, B}) - (B_{p, N} \cdot N_{1, A} + A_{p, N} \cdot N_{1, B})/N. \quad \dots\dots (10)$$

The next step is to calculate  $L_2$  and  $\lambda'_2$  from the equations

$$L_2 \sin \lambda'_2 = U_2/Kd_{m2S}. \quad \dots\dots\dots (11)$$

$$L_2 \cos \lambda'_2 = V_2/Kd_{m2S}. \quad \dots\dots\dots (12)$$

where  $K=0.4943 \text{ SN}\{1 - [(N_{1, A})^2 + (N_{1, B})^2]/N^2\}. \quad \dots\dots\dots (13)$

$$d_{m p S} = 1/(S) \sin \pi(m-p)\{\cot [\pi(m-p)/S] + \cot (\pi p/S)\} \quad \dots\dots\dots (14)$$

taking  $m=2(1-1/M)$  where  $M=29.5306$ , the number of mean solar days in the mean synodic period of the moon. Thus the values of  $d_{m2S}$  and  $1/d_{m2S}$ , for  $S=24, 12$ , and  $8$  respectively (corresponding to hourly, bi-hourly or tri-hourly data), are as follows:

TABLE 1  
VALUES OF  $d_{m2S}$  AND OF  $1/d_{m2S}$

$S=$	24	12	8
$d_{m2S}=$	0.9596	0.9619	0.9658
$1/d_{m2S}=$	1.0421	1.0396	1.0354

Finally a phase correction must be applied to the angle  $\lambda'_2$ , to give the true phase angle  $\lambda_2$ . This phase correction (to be added to  $\lambda'_2$ ) is given in degrees by

$$\lambda_2 - \lambda'_2 = 2L^\circ/M - 15^\circ mH' + m(L - L')^\circ. \quad \dots\dots\dots (15)$$

where  $L$  denotes the longitude measured from Greenwich (in degrees, reckoned positive if westward, negative if eastward, up to  $180^\circ$ ) of the station  $P$  to which the data refer;  $L'$  is the longitude (similarly measured and reckoned) of the meridian of time-reckoning with respect to which the data are tabulated;  $H'$  is the solar hour of the initial value ( $s=0$ ) of each daily sequence, according to the same time-reckoning;  $m=2(1-1/M)$ ,  $M$  being equal to  $29.5306$  mean solar days.

If the data are not instantaneous values but means for hourly intervals, strictly speaking there is also an amplitude correction-factor applicable to  $L_2$ , making allowance for the use of means; practically it can be neglected owing to the insignificance of this factor, being equal to  $1.0107$  in the present case.

Then the lunar semi-diurnal variation at the station  $P$  as thus determined is expressed by

$$L_2 \sin (2\tau + \lambda_2) \dots \dots \dots (16)$$

where  $\tau$  denotes mean lunar time reckoned in angle at the rate  $360^\circ$  or  $2\pi$  per mean lunar day, from the local upper or lower mean lunar transit across the meridian of  $P$ . The unit in which  $L_2$  so calculated is measured is the same as the unit used in the tabulated data, and at this stage it is easy to express the result in terms of any other chosen unit.

#### 8b. THE FINAL REDUCTIONS: GEOMAGNETIC AND SIMILAR DATA CONTAINING LUNAR SEMI-DIURNAL AND LUNI-SOLAR COMPONENTS

In this case, being the results described in Sections 6 (b), 6 (c), and 7, the numbers  $U_p$  and  $V_p$  (Section 8a) are calculated for  $p = 1, 2, 3, 4$ . Then the numbers  $L_n$  and  $\lambda'_n$  ( $n=1, 2, 3, 4$ ) are calculated from the formulae

$$L_n \sin \lambda'_n = \sum_{p=1}^4 (D_{mpS}/K) U_p \dots \dots \dots (17)$$

$$L_n \cos \lambda'_n = \sum_{p=1}^4 (D_{mpS}/K) V_p \dots \dots \dots (18)$$

where  $K$  is as defined above (13) and the factors  $D_{mpS}$  have the following values for the case of bi-hourly data:

TABLE 2

$D_{mpS}^\dagger, m=n-2/M, n=1, 2, 3, 4; S=12$				
$n$	$p=1$	2	3	4
1	1.0720	0.1445	0.1163	0.1112
2	-0.0360	1.0286	0.0994	0.0710
3	-0.0098	-0.0412	1.0150	0.0871
4	-0.0061	-0.0184	-0.0556	1.0114

† The factors  $D_{mpS}$  are expressible in terms of the determinant whose elements are  $(-1)^{n-p} d_{mpS}$  (equation 14).

The phases  $\lambda'_n$  must next be converted into  $\lambda_n$  by addition of the following corrections:

$$\lambda_n - \lambda'_n = 2L^\circ/M - 15^\circ(n-2/M)H' + (n-2/M)(L-L'). \dots (19)$$

The lunar and luni-solar harmonic components are thus determined in the form (6, Chap. VIII)

$$\sum_{n=1}^4 L_n \sin [2\tau + (n-2)t + \lambda_n], \dots \dots \dots (20)$$

where  $t$  denotes mean solar time reckoned in angle at the rate  $360^\circ$  or  $2\pi$  per mean solar day, from the epoch of local mean midnight at  $P$ .

#### 8c. IONOSPHERIC LUNAR AND LUNI-SOLAR SEMI-DIURNAL VARIATIONS

Investigations by Appleton and Weekes(4) for the  $E$  layer and by Martyn(7) for the  $F$  layer have already shown the existence of the lunar semi-diurnal variation of the type

$$L_2 \sin (2\tau + \lambda_2) \dots \dots \dots (21)$$



which is just the same as for some meteorological elements. Accordingly the determination of this harmonic from the necessary ionospheric data will be just the same as described in Section 8a.

However, Martyn's discussion shows that in addition there is a luni-solar semi-diurnal variation for the  $F_2$  layer which can be expressed\* in the following way

$$L \sin (4t-2\tau+\lambda) \dots\dots\dots (22)$$

It can also be readily shown\* that such a harmonic can be computed by a slight extension of the final stage in the computation of the harmonic  $L_2 \sin (2\tau+\lambda_2)$ .

(i) Instead of  $U_2$  and  $V_2$  we use  $U'_2$  and  $V'_2$ , which are calculated as follows (i.e. a mere change of some signs):

$$U'_2 = (A_{2,A} + B_{2,B}) - (A_{2,N} \cdot N_{1,A} + B_{2,N} \cdot N_{1,B})/N. \dots\dots (23)$$

$$V'_2 = (B_{2,A} - A_{2,B}) - (B_{2,N} \cdot N_{1,A} - A_{2,N} \cdot N_{1,B})/N. \dots\dots (24)$$

(ii) The factor  $d_{m2S}$ , calculated from the expression (14), taking  $m=2(1+1/M)$ , instead of  $m=2(1-1/M)$ , is 1.0229 for the case of bi-hourly data.

(iii) Insert  $m=2(1+1/M)$ , instead of  $m=2(1-1/M)$ , in the expression giving the phase correction.

In this way the lunar and luni-solar semi-diurnal variations for the  $F_2$  layer are expressed in the form

$$L_2 \sin (2\tau+\lambda_2) + L \sin (4t-2\tau+\lambda). \dots\dots\dots (25)$$

## 9. THE SOLAR DAILY HARMONIC COMPONENTS

The first four harmonic components of the solar daily variation are formed as follows: Using the numbers  $A_{p,N}$ ,  $B_{p,N}$  determined as in Section 6 (c) we calculate  $S_p$  and  $\sigma'_p$  from the equations

$$S_p \sin \sigma'_p = 2A_{p,N}/SN \dots\dots\dots (26)$$

$$S_p \cos \sigma'_p = 2B_{p,N}/SN \dots\dots\dots (27)$$

for  $p=1, 2, 3, 4$ . Phase corrections as follows are then added to  $\sigma'_p$  to obtain the true phases  $\sigma_p$

$$\sigma_p - \sigma'_p = -p(15^\circ H' - L + L') \dots\dots\dots (28)$$

The solar daily variation is then expressed (up to the first four components) by

$$\sum_{p=1}^4 S_p \sin (pt + \sigma_p) \dots\dots\dots (29)$$

## 10. DETERMINATION OF PROBABLE ERRORS

The probable error for each lunar harmonic  $p$  (or  $n$ ) is then determined from the 12 pairs of numbers  $A_{p,r}$  and  $B_{p,r}$  (Section 6) as follows: Divide each pair  $A_{p,r}$  and  $B_{p,r}$  by its corresponding  $N_r$  ( $r=0$  to  $r=11$ ), getting 12 pairs  $A_{p,r}/N_r$  and  $B_{p,r}/N_r$ . Calculate the mean value of the 12 numbers  $A_{p,r}/N_r$ , and subtract it from each, and likewise for the 12 numbers  $B_{p,r}/N_r$ ; this process gives 12 further pairs of numbers which can be denoted by  $\triangle a_{p,r}$  and  $\triangle b_{p,r}$

\* The proof is here omitted.

where  $r=0$  to  $r=11$  as before. Now evaluate the 12 quantities  $\sqrt{(\Delta a_{p,r})^2 + (\Delta b_{p,r})^2}$  and their mean value  $R^*$ . Also calculate  $\sqrt{(U_p)^2 + (V_p)^2}/N = L^*$  as from Section 8, and form the ratio  $L^*/R^*$ . Then the probable error (P.E.) of the mean determination of the  $p$ -th lunar harmonic can be easily found by multiplying  $R^*$  by the corresponding ratio P.E./ $R^*$  given in the following table.

TABLE 3  
PROBABLE ERRORS FOR BI-HOURLY DATA, FIRST METHOD

$L^*/R^*$	P.E./ $R^*$	$L^*/R^*$	P.E./ $R^*$
0.0	0.0512	0.4	0.0478
0.05	0.0511	0.45	0.0469
0.1	0.0510	0.5	0.0457
0.15	0.0507	0.55	0.0445
0.2	0.0504	0.6	0.0430
0.25	0.0499	0.65	0.0413
0.3	0.0493	0.7	0.0393
0.35	0.0486	0.75	0.0370

When the ratio  $L^*/R^*$  exceeds about 0.75, a more refined procedure must be followed. Evaluate  $\Delta a'_{p,r}$  and  $\Delta b'_{p,r}$  from  $\Delta a_{p,r}$  and  $\Delta b_{p,r}$  according to the formulae

$$\Delta a'_{p,r} = A_r \Delta a_{p,r} + B_r \Delta b_{p,r} \quad \dots \dots \dots (30)$$

$$\Delta b'_{p,r} = C_r \Delta a_{p,r} + D_r \Delta b_{p,r} \quad \dots \dots \dots (31)$$

where  $A_r, B_r, C_r, D_r$  are simple multiplying factors as given by the following table.

TABLE 4  
VALUES OF FACTORS  $A_r, B_r, C_r, D_r$

$r$	$A_r$	$B_r$	$C_r$	$D_r$
0	1	0	0	1
1	0.866	0.5	-0.5	0.866
2	0.5	0.866	-0.866	0.5
3	0	1	-1	0
4	-0.5	0.866	-0.866	-0.5
5	-0.866	0.5	-0.5	-0.866
6	-1	0	0	-1
7	-0.866	-0.5	0.5	-0.866
8	-0.5	-0.866	0.866	-0.5
9	0	-1	1	0
10	0.5	-0.866	0.866	0.5
11	0.866	-0.5	0.5	0.866

Next calculate the mean value of the 12 numbers  $\Delta a'_{p,r}$ , and subtract it from each, and likewise for the 12 numbers  $\Delta b'_{p,r}$ ; this process gives 12 further

pairs of numbers which can be denoted by  $\Delta a''_{p, r}$  and  $\Delta b''_{p, r}$ . As before we evaluate 12 quantities  $\sqrt{(\Delta a''_{p, r})^2 + (\Delta b''_{p, r})^2}$  and their mean value  $R'$ , and thence form the ratio  $L^*/R'$ ; the probable error (P.E.) of the mean determination of the  $p$ -th lunar harmonic is then found by multiplying  $R'$  by the corresponding ratio P.E./ $R'$  given in the following table.

TABLE 5  
PROBABLE ERROR FOR BI-HOURLY DATA, SECOND METHOD

$L^*/R'$	P.E./ $R'$	$L^*/R'$	P.E./ $R'$
0	0.0512	6	0.0491
1	0.0511	7	0.0483
2	0.0510	8	0.0473
3	0.0507	9	0.0463
4	0.0503	10	0.0450
5	0.0497	11	0.0436
6	0.0491	12	0.0420

The probable error is used to indicate the reliability of the corresponding lunar harmonic thus determined. The smaller the probable error as compared with the magnitude of the corresponding lunar harmonic, the better is the determination. Generally speaking any determination of the lunar harmonic will not be considered as reliable if the probable error exceeds one-third of the amplitude of the lunar harmonic.

The probable errors as determined above can also be applied to the corresponding solar harmonics; but they are not of special interest in this case because the ratios of the amplitudes to the corresponding P.E.'s are usually so large.

## 11. ARRANGEMENT OF THE REDUCTIONS

The work of computation from Section 6 onwards (the twelve group-sum sequences and the total-sum sequence of Section 5 having been calculated) can be lightened by suitable arrangement of the computing form used. This will be illustrated here in some detail for the case  $S=12$  (bi-hourly data); the cases  $S=24$  (hourly data), and  $S=8$  (tri-hourly data) can be similarly treated.

### 11. (a) Schemes for Harmonic Analyses

In the cases as described in Sections 6 (b), 6 (c) and 9, where the numbers  $A_{p, r}$  and  $B_{p, r}$ , or  $A_{p, N}$  and  $B_{p, N}$ , are to be calculated for  $p=1, 2, 3, 4$  from their respective sum-sequences  $g_{s, r}$  or  $g_{s, N}$ , we use the following scheme (here dropping the suffix  $r$  or  $N$  for brevity): Divide each sequence of bi-hourly sums  $g_0$  to  $g_{11}$  into half-sequences to be added and subtracted in rows, the non-cyclic correction  $d=g_{12}-g_0$  being always included in the first column of both rows, i.e.



SCHEME A

	$g_{12}$		$g_0$	$g_1$	$g_2$	$g_3$	$g_4$	$g_5$
	$g_0$		$g_6$	$g_7$	$g_8$	$g_9$	$g_{10}$	$g_{11}$
d and sums	$d$		$s_0$	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$
Diffs.	$d$		$d_0$	$d_1$	$d_2$	$d_3$	$d_4$	$d_5$

Clearly here the suffix  $s$  attached to  $g_{i,r}$  or  $g_{i,x}$  denotes the series-number of the sum-sequence considered. (It must not be confused with the usual hourly order.) For  $p=1$  and  $p=3$ , we use the differences, whereas for  $p=2$  and  $p=4$ , we use the sums together with the non-cyclic correction  $d$ . Then the formulae (3)-(4) of Section 6 become

$$A_1 = d_0 + 0.866(d_1 - d_5) + 0.5(d_2 - d_4 + d) \dots\dots\dots (32)$$

$$B_1 = 0.5(d_1 + d_5) + 0.866(d_2 + d_4) + d_3 + 1.866d \dots\dots\dots (33)$$

$$A_3 = d_0 - d_2 + d_4 + 0.5d \dots\dots\dots (34)$$

$$B_3 = d_1 - d_3 + d_5 + 0.5d \dots\dots\dots (35)$$

$$A_2 = s_0 - s_3 + 0.5(s_1 - s_2 - s_4 + s_5 + d) = 2A'_2 \dots\dots\dots (36)$$

$$B_2 = 0.866(s_1 + s_2 - s_4 - s_5 + d) = 0.866B'_2 \dots\dots\dots (37)$$

$$A_4 = s_0 + s_3 - 0.5(s_1 + s_2 + s_4 + s_5 + d) = 2A'_4 \dots\dots\dots (38)$$

$$B_4 = 0.866(s_1 - s_2 + s_4 - s_5 - d/3) = 0.866B'_4 \dots\dots\dots (39)$$

where the suffix  $r$  or  $N$  may be added to the  $A$ 's and  $B$ 's, in accordance with the  $g$ 's. Thus the calculation of the numbers  $A_{i,r}$  and  $B_{i,r}$ , or  $A_{i,x}$  and  $B_{i,x}$  ( $p=1, 2, 3, 4$ ) becomes quick and easy, involving only very simple multiplications followed by additions.

In the cases of Section 6 (*a*) where the numbers  $A_{i,r}$ ,  $B_{i,r}$  are calculated only for  $p=2$  from their corresponding ( $r$ -th) group sum sequences, we use the following scheme for each sum-sequence,

SCHEME B

	$g_{12}$		$g_0$	$g_1$	$g_2$	$g_3$	$g_4$	$g_5$
	$g_0$		$g_6$	$g_7$	$g_8$	$g_9$	$g_{10}$	$g_{11}$
d and sums	$d$		$s_0$	$s_1$	$s_2$	$s_3$	$s_4$	$s_5$

and the formulae (36)-(37) are then duly applied.

Similarly, the secondary harmonic analyses (Section 7) of the 12-term sequences  $N_r$ ,  $A_{i,r}$ ,  $B_{i,r}$  according to the formulae (6)-(8) can also be simplified by first dividing each sequence of  $N_r$ ,  $A_{i,r}$ , and  $B_{i,r}$  ( $r=0$  to  $r=11$ ) into two half-sequences to be subtracted, the numerical factors by which the differences (here denoted by  $\frac{1}{2}N_r$ ,  $\frac{1}{2}A_{i,r}$ , and  $\frac{1}{2}B_{i,r}$  respectively) of these half-sequences must be multiplied being then limited to 0, 0.5, 0.866, and 1. As the three

sets of formulae (6)-(8) have the same form, we take the sequence  $N_r$  as an example, and indicate the scheme as follows :

SCHEME C

$r$	$\Delta N_r$	$\text{Cos } \pi r/6 \cdot \Delta N_r$	$\text{Sin } \pi r/6 \cdot \Delta N_r$
0	$N_0-N_6$	$\Delta N_0$	0
1	$N_1-N_7$	$0.866 \Delta N_1$	$0.5 \Delta N_1$
2	$N_2-N_8$	$0.5 \Delta N_2$	$0.866 \Delta N_2$
3	$N_3-N_9$	0	$\Delta N_3$
4	$N_4-N_{10}$	$-0.5 \Delta N_4$	$0.866 \Delta N_4$
5	$N_5-N_{11}$	$-0.866 \Delta N_5$	$0.5 \Delta N_5$
		$\Sigma \text{ gives } N_{1, A}$	$\Sigma \text{ gives } N_{1, B}$

### 11. (b) Computing Form for the Lunar Semi-diurnal Term Only

Figure 1 shows a suggested arrangement of a computing form with a worked example of the determination of air-tide for the  $J$  season (May-August) of the period 1924-1938 inclusive at the Huancayo Magnetic Observatory, Peru. For convenience of printing, Figure 1 shows the arrangement of the form on reduced scale and Figures 2-7 show the separate parts of the form with the actual computations for this case. Also in the Figures 2-7, the words, symbols, and numbers that are here italicized, together with some necessary lines (vertical and horizontal), are supposed to have been printed on the computing sheet ; all the rest refer to the present worked example. The computing form is devised in such a way that it includes all necessary computing steps other than the summing of the individual daily sequences (Section 4). It can also be noted that the separate parts of the form are closely related.

#### I. Description

II. The 12 lunar-age group-sum sequences and their primary harmonic analyses (Sections 5, 6, 6 (a), and 11 (a))	IV. The secondary harmonic analyses and final reduction (Sections 7, 8a, and 11 (a))
	V. The probable - error calculation (Section 10)
III. The total-sum sequence and its harmonic analysis (Sections 5, 6 (c), 9, and 11 (a))	VI. The solar daily harmonic components (Section 9)

Fig. 1.—The lay-out of the computing form.

The form of Section I of Figure 1 is shown in Figure 2 ; it includes all necessary particulars, namely, the name of the station, its geographical co-ordinates (i.e. latitude, longitude, and height), the geophysical element considered,

the period of the data, the set (Section 4), the working unit (i.e. the unit of the last digit used in the tabulations), the time-reckoning used, the initial hour ( $H'$ ) of the daily sequence, and other remarks such as changes in the instruments used or their locations, and so on. Figure 2 is here shown to consist of five rows: actually the number of rows may be reduced by printing all the item-names on one row, and writing the corresponding particulars immediately below them.

The form of Section II is illustrated in Figure 3; it contains twelve lunar-age ( $r=0$  to 11) group-sum sequences and their primary analyses, together with the number of daily-sequences ( $N_r$ ) in each group. For each lunar-age group ( $r$ ), the group-sum sequence ( $g_{s,r}$ ) is entered from the group-summation (Section 5) according to Scheme *B* of Section 11 (*a*); then the difference  $d$  and the six sums  $s_0 \dots s_5$  are subsequently made, the former being preferably in blue ink, but here enclosed in round brackets, the latter being preferably in red ink, but here enclosed in square brackets. Then  $A'_{2,r}$  is calculated, for each lunar-age group, by multiplying the bracketed numbers for that group by the upper factors, printed alongside each, on the right, and adding the products; similarly for  $B'_{2,r}$ , using the lower printed factors instead of the upper. These  $A'_{2,r}$  and  $B'_{2,r}$  ( $r=0$  to  $r=11$ ) are shown in the last column but one in Section II. In the last column are shown the numbers  $A_{2,r}$  and  $B_{2,r}$  obtained by multiplying  $A'_{2,r}$  and  $B'_{2,r}$  by the respective factors 0.5 and 0.866.

The form of Section III is illustrated in Figure 4; it contains one total-sum sequence ( $g_{s,x}$ ) and its harmonic analyses, together with the corresponding number  $N$ . The entry of the total-sum sequence from the total-summation (Section 5) and the subsequent calculation of the numbers  $A_{p,x}$  and  $B_{p,x}$  are as described for Section II except that (1) Scheme *A* of Section 11 (*a*) is used, and (2)  $A_{p,x}$  and  $B_{p,x}$  ( $p=1, 2, 3, 4$ ) should be calculated for each value of  $p$ .

Figure 5 shows Section IV, which includes the secondary harmonic analyses and the final reductions. The 12-term sequences  $A_{2,r}$ ,  $B_{2,r}$ , and  $N_r$  of Section II are analysed according to Scheme *C* of Section 11 (*a*). The final reductions follow as described in Section 8 (*a*).

Section V (Figure 6) is devoted to the determination (Section 10) of the probable error of the semi-diurnal lunar result derived in Section IV;  $R^*$  is first calculated, and if  $L^* R^*$  is 0.75 or less (Section 10), the remaining columns of this section leading to  $R'$ , are not computed. In the example shown, however, the calculation of  $R'$  is necessary.

---

*Determination of lunar and solar daily variations*

---

<i>Station</i> : Huancayo, Peru	<i>Element</i> : Air pressure	<i>Time-reckoning</i> : 75° W. standard
<i>Latitude</i> : 12° 3' S.	<i>Period</i> : 1924-1938	<i>Initial hour</i> : 1st hour
<i>Longitude</i> : 75° 20' W.	<i>Set</i> : J (May-Aug.)	<i>Other remarks</i> :
<i>Height</i> : 11,000 ft.	<i>Working unit</i> : 0.1 mm. Hg	

---

Fig. 2.—Section I of the computing form (Section 11 (*b*)).

Section VI (Figure 7) concerns the determination (Section 9) of the solar daily harmonic components, from the results in the last column of Section III.



	N	$g_{12}, r$ $g_{01}, r$	$g_{10}, r$ $g_{61}, r$	$g_{11}, r$ $g_{71}, r$	$g_{21}, r$ $g_{81}, r$	$g_{31}, r$ $g_{91}, r$	$g_{41}, r$ $g_{101}, r$	$g_{51}, r$ $g_{111}, r$	$A', r$ $B', r$	$A_{21}, r$ $B_{21}, r$
0	156	26575 26648 (-73)	26648 25458 [52106] I I	26016 23069 2 [49085] 0	25964 22718 I [48682] I	26945 24520 -I [51465] I	28343 26018 -2 [54361] 0	27738 26653 -I [54391] -I	I [1642] -I [-11058]	1/2 821.0 1/2 821.0 $\sqrt{3/2}$ -9576.6
1	154	26373 26383 (-10)	26383 25416 [51799] I I	25739 23101 2 [48840] 0	25718 22757 I [48475] I	26759 24403 -I [51162] I	28175 25846 -2 [54021] 0	27643 26522 -I [54165] -I	I [1773] -I [-10881]	1/2 886.5 1/2 886.5 $\sqrt{3/2}$ -9423.3
2	153	26225 26082 (143)	26082 25016 [51098] I I	25446 22703 2 [48149] 0	25475 22459 I [47934] I	26527 24257 -I [50784] I	27918 25788 -2 [53706] 0	27377 26355 -I [53632] -I	I [912] -I [-11112]	1/2 456.0 1/2 456.0 $\sqrt{3/2}$ -9623.3
3	152	25848 25960 (-112)	25960 24709 [50669] I I	25344 22367 2 [47711] 0	25435 22110 I [47545] I	26499 23953 -I [50452] I	27771 25463 -2 [53234] 0	27035 25996 -I [53031] -I	I [285] -I [-11121]	1/2 142.5 1/2 142.5 $\sqrt{3/2}$ -9631.1
4	150	25657 25703 (-46)	25703 24544 [50247] I I	25101 22245 2 [47346] 0	25145 22056 I [47201] I	26153 23905 -I [50058] I	27521 25402 -2 [52923] 0	26810 25899 -I [52709] -I	I [263] -I [-11131]	1/2 131.5 1/2 131.5 $\sqrt{3/2}$ -9639.8
5	150	25755 25681 (74)	25681 24631 [50312] I I	25139 22352 2 [47491] 0	25219 22160 I [47379] I	26268 23934 -I [50202] I	27569 25364 -2 [52933] 0	26870 25891 -I [52761] -I	I [234] -I [-10750]	1/2 117.0 1/2 117.0 $\sqrt{3/2}$ -9309.8

Fig. 3.—Section II of the computing form (Section 11(b)).

$\lambda$	$Q_{10}$	$Q_{11}$	$Q_{12}$	$Q_{13}$	$Q_{14}$	$Q_{15}$	$Q_{16}$	$Q_{17}$	$Q_{18}$	$Q_{19}$	$Q_{20}$	$Q_{21}$	$Q_{22}$	$Q_{23}$	$Q_{24}$	$Q_{25}$	$Q_{26}$	$Q_{27}$	$Q_{28}$	$Q_{29}$	$Q_{30}$	$Q_{31}$	$Q_{32}$	$Q_{33}$	$Q_{34}$	$Q_{35}$	$Q_{36}$	$Q_{37}$	$Q_{38}$	$Q_{39}$	$Q_{40}$	$Q_{41}$	$Q_{42}$	$Q_{43}$	$Q_{44}$	$Q_{45}$	$Q_{46}$	$Q_{47}$	$Q_{48}$	$Q_{49}$	$Q_{50}$	$Q_{51}$	$Q_{52}$	$Q_{53}$	$Q_{54}$	$Q_{55}$	$Q_{56}$	$Q_{57}$	$Q_{58}$	$Q_{59}$	$Q_{60}$	$Q_{61}$	$Q_{62}$	$Q_{63}$	$Q_{64}$	$Q_{65}$	$Q_{66}$	$Q_{67}$	$Q_{68}$	$Q_{69}$	$Q_{70}$	$Q_{71}$	$Q_{72}$	$Q_{73}$	$Q_{74}$	$Q_{75}$	$Q_{76}$	$Q_{77}$	$Q_{78}$	$Q_{79}$	$Q_{80}$	$Q_{81}$	$Q_{82}$	$Q_{83}$	$Q_{84}$	$Q_{85}$	$Q_{86}$	$Q_{87}$	$Q_{88}$	$Q_{89}$	$Q_{90}$	$Q_{91}$	$Q_{92}$	$Q_{93}$	$Q_{94}$	$Q_{95}$	$Q_{96}$	$Q_{97}$	$Q_{98}$	$Q_{99}$	$Q_{100}$	$Q_{101}$	$Q_{102}$	$Q_{103}$	$Q_{104}$	$Q_{105}$	$Q_{106}$	$Q_{107}$	$Q_{108}$	$Q_{109}$	$Q_{110}$	$Q_{111}$	$Q_{112}$	$Q_{113}$	$Q_{114}$	$Q_{115}$	$Q_{116}$	$Q_{117}$	$Q_{118}$	$Q_{119}$	$Q_{120}$	$Q_{121}$	$Q_{122}$	$Q_{123}$	$Q_{124}$	$Q_{125}$	$Q_{126}$	$Q_{127}$	$Q_{128}$	$Q_{129}$	$Q_{130}$	$Q_{131}$	$Q_{132}$	$Q_{133}$	$Q_{134}$	$Q_{135}$	$Q_{136}$	$Q_{137}$	$Q_{138}$	$Q_{139}$	$Q_{140}$	$Q_{141}$	$Q_{142}$	$Q_{143}$	$Q_{144}$	$Q_{145}$	$Q_{146}$	$Q_{147}$	$Q_{148}$	$Q_{149}$	$Q_{150}$	$Q_{151}$	$Q_{152}$	$Q_{153}$	$Q_{154}$	$Q_{155}$	$Q_{156}$	$Q_{157}$	$Q_{158}$	$Q_{159}$	$Q_{160}$	$Q_{161}$	$Q_{162}$	$Q_{163}$	$Q_{164}$	$Q_{165}$	$Q_{166}$	$Q_{167}$	$Q_{168}$	$Q_{169}$	$Q_{170}$	$Q_{171}$	$Q_{172}$	$Q_{173}$	$Q_{174}$	$Q_{175}$	$Q_{176}$	$Q_{177}$	$Q_{178}$	$Q_{179}$	$Q_{180}$	$Q_{181}$	$Q_{182}$	$Q_{183}$	$Q_{184}$	$Q_{185}$	$Q_{186}$	$Q_{187}$	$Q_{188}$	$Q_{189}$	$Q_{190}$	$Q_{191}$	$Q_{192}$	$Q_{193}$	$Q_{194}$	$Q_{195}$	$Q_{196}$	$Q_{197}$	$Q_{198}$	$Q_{199}$	$Q_{200}$	$Q_{201}$	$Q_{202}$	$Q_{203}$	$Q_{204}$	$Q_{205}$	$Q_{206}$	$Q_{207}$	$Q_{208}$	$Q_{209}$	$Q_{210}$	$Q_{211}$	$Q_{212}$	$Q_{213}$	$Q_{214}$	$Q_{215}$	$Q_{216}$	$Q_{217}$	$Q_{218}$	$Q_{219}$	$Q_{220}$	$Q_{221}$	$Q_{222}$	$Q_{223}$	$Q_{224}$	$Q_{225}$	$Q_{226}$	$Q_{227}$	$Q_{228}$	$Q_{229}$	$Q_{230}$	$Q_{231}$	$Q_{232}$	$Q_{233}$	$Q_{234}$	$Q_{235}$	$Q_{236}$	$Q_{237}$	$Q_{238}$	$Q_{239}$	$Q_{240}$	$Q_{241}$	$Q_{242}$	$Q_{243}$	$Q_{244}$	$Q_{245}$	$Q_{246}$	$Q_{247}$	$Q_{248}$	$Q_{249}$	$Q_{250}$	$Q_{251}$	$Q_{252}$	$Q_{253}$	$Q_{254}$	$Q_{255}$	$Q_{256}$	$Q_{257}$	$Q_{258}$	$Q_{259}$	$Q_{260}$	$Q_{261}$	$Q_{262}$	$Q_{263}$	$Q_{264}$	$Q_{265}$	$Q_{266}$	$Q_{267}$	$Q_{268}$	$Q_{269}$	$Q_{270}$	$Q_{271}$	$Q_{272}$	$Q_{273}$	$Q_{274}$	$Q_{275}$	$Q_{276}$	$Q_{277}$	$Q_{278}$	$Q_{279}$	$Q_{280}$	$Q_{281}$	$Q_{282}$	$Q_{283}$	$Q_{284}$	$Q_{285}$	$Q_{286}$	$Q_{287}$	$Q_{288}$	$Q_{289}$	$Q_{290}$	$Q_{291}$	$Q_{292}$	$Q_{293}$	$Q_{294}$	$Q_{295}$	$Q_{296}$	$Q_{297}$	$Q_{298}$	$Q_{299}$	$Q_{300}$	$Q_{301}$	$Q_{302}$	$Q_{303}$	$Q_{304}$	$Q_{305}$	$Q_{306}$	$Q_{307}$	$Q_{308}$	$Q_{309}$	$Q_{310}$	$Q_{311}$	$Q_{312}$	$Q_{313}$	$Q_{314}$	$Q_{315}$	$Q_{316}$	$Q_{317}$	$Q_{318}$	$Q_{319}$	$Q_{320}$	$Q_{321}$	$Q_{322}$	$Q_{323}$	$Q_{324}$	$Q_{325}$	$Q_{326}$	$Q_{327}$	$Q_{328}$	$Q_{329}$	$Q_{330}$	$Q_{331}$	$Q_{332}$	$Q_{333}$	$Q_{334}$	$Q_{335}$	$Q_{336}$	$Q_{337}$	$Q_{338}$	$Q_{339}$	$Q_{340}$	$Q_{341}$	$Q_{342}$	$Q_{343}$	$Q_{344}$	$Q_{345}$	$Q_{346}$	$Q_{347}$	$Q_{348}$	$Q_{349}$	$Q_{350}$	$Q_{351}$	$Q_{352}$	$Q_{353}$	$Q_{354}$	$Q_{355}$	$Q_{356}$	$Q_{357}$	$Q_{358}$	$Q_{359}$	$Q_{360}$	$Q_{361}$	$Q_{362}$	$Q_{363}$	$Q_{364}$	$Q_{365}$	$Q_{366}$	$Q_{367}$	$Q_{368}$	$Q_{369}$	$Q_{370}$	$Q_{371}$	$Q_{372}$	$Q_{373}$	$Q_{374}$	$Q_{375}$	$Q_{376}$	$Q_{377}$	$Q_{378}$	$Q_{379}$	$Q_{380}$	$Q_{381}$	$Q_{382}$	$Q_{383}$	$Q_{384}$	$Q_{385}$	$Q_{386}$	$Q_{387}$	$Q_{388}$	$Q_{389}$	$Q_{390}$	$Q_{391}$	$Q_{392}$	$Q_{393}$	$Q_{394}$	$Q_{395}$	$Q_{396}$	$Q_{397}$	$Q_{398}$	$Q_{399}$	$Q_{400}$	$Q_{401}$	$Q_{402}$	$Q_{403}$	$Q_{404}$	$Q_{405}$	$Q_{406}$	$Q_{407}$	$Q_{408}$	$Q_{409}$	$Q_{410}$	$Q_{411}$	$Q_{412}$	$Q_{413}$	$Q_{414}$	$Q_{415}$	$Q_{416}$	$Q_{417}$	$Q_{418}$	$Q_{419}$	$Q_{420}$	$Q_{421}$	$Q_{422}$	$Q_{423}$	$Q_{424}$	$Q_{425}$	$Q_{426}$	$Q_{427}$	$Q_{428}$	$Q_{429}$	$Q_{430}$	$Q_{431}$	$Q_{432}$	$Q_{433}$	$Q_{434}$	$Q_{435}$	$Q_{436}$	$Q_{437}$	$Q_{438}$	$Q_{439}$	$Q_{440}$	$Q_{441}$	$Q_{442}$	$Q_{443}$	$Q_{444}$	$Q_{445}$	$Q_{446}$	$Q_{447}$	$Q_{448}$	$Q_{449}$	$Q_{450}$	$Q_{451}$	$Q_{452}$	$Q_{453}$	$Q_{454}$	$Q_{455}$	$Q_{456}$	$Q_{457}$	$Q_{458}$	$Q_{459}$	$Q_{460}$	$Q_{461}$	$Q_{462}$	$Q_{463}$	$Q_{464}$	$Q_{465}$	$Q_{466}$	$Q_{467}$	$Q_{468}$	$Q_{469}$	$Q_{470}$	$Q_{471}$	$Q_{472}$	$Q_{473}$	$Q_{474}$	$Q_{475}$	$Q_{476}$	$Q_{477}$	$Q_{478}$	$Q_{479}$	$Q_{480}$	$Q_{481}$	$Q_{482}$	$Q_{483}$	$Q_{484}$	$Q_{485}$	$Q_{486}$	$Q_{487}$	$Q_{488}$	$Q_{489}$	$Q_{490}$	$Q_{491}$	$Q_{492}$	$Q_{493}$	$Q_{494}$	$Q_{495}$	$Q_{496}$	$Q_{497}$	$Q_{498}$	$Q_{499}$	$Q_{500}$
-----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------	-----------

$r$	$N$	$g_{12}, N$	$g_{60}, N$	$g_{70}, N$	$g_{20}, N$	$g_{20}, N$	$g_{40}, N$	$g_{110}, N$	$A', P', N$ $B', P', N$ $p = \frac{2}{3}, \frac{4}{3}, 1$	$A', P', N$ $B', P', N$ $p = \frac{2}{3}, \frac{4}{3}, 1$
$AM$	1833	313229	313146	306307	306712	318859	334729	326919		
		313146	300655	272348	269334	290513	308244	314853		
(83)	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
(83)	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$
	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$

Fig. 4.—Section III of the computing form (Section 11(b)).



Such computations are to be made for each set (Section 4) on a separate computing sheet such as Figure 1. Evidently the same procedure may be followed for the whole collection of data. Preferably in practice, however, some omissions may be made. Consider the sheets for three (e.g. *J*, *E*, *D* seasons) or more sets into which the whole material is divided. By adding corresponding figures in the second and last columns of Section II (Figure 3) for these sheets, a "total" sheet (like Figure 1) can be obtained, with only the second and last columns of Section II filled in, the part showing twelve lunar-age group-sum sequences being here unnecessary. Then the analyses in Sections III, IV, and VI can be done on this "total" sheet as before; this will afford not only the required harmonics for the whole collection of data, but also some useful checks as described later in Section 11 (*e*). The P.E. for this "total" sheet is usually determined as follows: Form the mean value (weighted or otherwise) of the P.E.'s for all the sets in the collection considered, and divide it by the square root of the number of sets involved.

$\Delta A_2, \tau$	$\Delta A_2, \tau$ $\cos \frac{\pi \tau}{6}$	$\Delta A_2, \tau$ $\sin \frac{\pi \tau}{6}$	$\Delta B_2, \tau$	$\Delta B_2, \tau$ $\cos \frac{\pi \tau}{6}$	$\Delta B_2, \tau$ $\sin \frac{\pi \tau}{6}$	$\Delta N_T$	$\Delta N_T$ $\cos \frac{\pi \tau}{6}$	$\Delta N_T$ $\sin \frac{\pi \tau}{6}$
713.0	713.0	x	129.9	129.9	x	0	0	x
745.0	645.2	372.5	-545.6	-472.5	-272.8	5	4.33	2.50
345.5	172.8	299.2	-522.2	-261.1	-452.2	0	0	0
-253.5	x	-253.5	-666.0	x	-666.0	0	x	0
-375.5	187.8	-325.2	-435.6	217.8	-377.2	-4	2.00	-3.46
-640.0	554.3	-320.0	181.9	-157.5	90.9	-4	3.46	-2.00
Sum	$\Delta_{2,A}$ 2273.1	$\Delta_{2,B}$ -227.0	Sum	$B_{2,A}$ -543.4	$B_{2,B}$ -1677.3	Sum	$\Sigma_{1,A}$ 9.79	$\Sigma_{1,B}$ -2.96

$$\begin{aligned}
 V_2 &= (A_2, A - B_2, B) - (A_2, N \cdot N_1, A - B_2, N \cdot N_1, B) \cdot N = 4107.8 \\
 V_3 &= (A_2, B + B_2, A) - (B_2, N \cdot N_1, A + A_2, N \cdot N_1, B) \cdot N = 161.9 \\
 K &= 5.9316 N \{1 - [(N_1, A)^2 + (N_1, B)^2] / N^2\} = 10872.2843 \\
 L_2 \sin \lambda'_2 &= 1.0396 U_2 / K = 0.39278 \quad \therefore L_2 = 0.3931 \text{ units} = 52.4 \text{ microbars} \\
 L_2 \cos \lambda'_2 &= 1.0396 V_2 / K = -0.01548 \quad \therefore \lambda'_2 = 92^\circ 25' \\
 \lambda_2 - \lambda'_2 &= -2(15H' - L - L') + \frac{2}{M}(L' + 15H') = -23^\circ 23' \quad \therefore \lambda_2 = 69.0^\circ
 \end{aligned}$$

Fig. 5. -Section IV of the computing form (Section 11(b)).

#### 11. (*c*) *Computing Form for the Determination of the Lunar Semi-diurnal and Luni-solar Components in Geomagnetic and Similar Data*

For such data the computing form for each set (e.g. *J* season) of data for a certain element (e.g. geomagnetic declination) will have, as compared with the form already fully described in Section 11 (*b*), the following characteristics or modifications:

(i) Sections 1, III, VI remain unchanged (Sections 5, 6 (*c*), 9, 11 (*a*), and 11 (*b*)).

(ii) Section II is modified and extended, so that for each lunar-age group, the group-sum sequences ( $g_s$ , ...) are entered from the group-summation (Section 5) and subsequently analysed in the manner already fully described for Section III.

This is because for such data (Section 6 (b)) the numbers  $A_{p, \tau}$  and  $B_{p, \tau}$  should be calculated for four values of  $p$  ( $p = 1, 2, 3, 4$ ), giving 96 numbers  $A_{p, \tau}$  and  $B_{p, \tau}$  (see Sections 5, 6, 6 (b), 11 (a), and 11 (b)).

(iii) Section IV is extended, because four calculations (Section 7) of  $A_{p, A}$ ,  $A_{p, B}$ ;  $B_{p, A}$ ,  $B_{p, B}$ ; and  $U_p$ ,  $V_p$  are made, one for each value of  $p$  ( $p = 1, 2, 3, 4$ ). The calculation of  $N_{1, A}$  and  $N_{1, B}$  is made as before. Then follow the calculations of  $L_n$ ,  $\lambda'_n$ , and  $\lambda_n$  for  $n = 1, 2, 3, 4$ , according to the formulae (17)-(19) of Section 8b (see Sections 7, 8b, 11 (a), and 11 (b)).

$A_{2, \tau}/N_{\tau}$	$B_{2, \tau}/N_{\tau}$	$\Delta a_2$	$\Delta b_2$	$\sqrt{(\Delta a_2)^2 + (\Delta b_2)^2}$	$\Delta a_2'$	$\Delta b_2'$	$\Delta a_2''$	$\Delta b_2''$	$\sqrt{(\Delta a_2'')^2 + (\Delta b_2'')^2}$
5.26	-61.39	2.76	0.01	2.76	2.76	0.01	2.56	-0.14	2.56
5.76	-61.19	3.26	0.21	3.27	2.93	-1.45	2.73	-1.60	3.16
2.98	-62.90	0.48	-1.50	1.57	-1.06	-1.17	-1.26	-1.32	1.83
0.94	-63.36	-1.56	-1.96	2.50	-1.96	1.56	-2.16	1.41	2.58
0.88	-64.27	-1.62	-2.87	3.30	-1.68	2.84	-1.88	2.69	3.28
0.78	-62.97	1.72	0.67	1.85	1.15	1.44	0.95	1.29	1.60
0.69	-62.22	-1.81	-0.82	1.99	1.81	0.82	1.61	0.67	1.74
0.95	-59.58	-1.55	1.82	2.39	0.43	-2.35	0.23	-2.50	2.51
0.72	-59.48	-1.78	1.92	2.62	-0.77	-2.50	-0.97	-2.65	2.82
2.60	-58.98	0.10	2.42	2.42	-2.42	0.10	-2.62	-0.05	2.62
3.29	-59.77	0.79	1.63	1.81	-1.02	1.50	-1.22	1.35	1.82
4.92	-61.63	2.42	-0.23	2.43	2.21	1.01	2.01	0.86	2.19
28.77	7.56.4			28.91	2.35	1.81			28.71
2.50	-61.40			$R^* = 2.41$	0.20	0.15			$R' = 2.39$

$$L^* = \sqrt{U^2 + V^2}/N = 2.243 \quad L^*/R^* = 0.931$$

$$L^*/R' = 0.938$$

From Table,  $P.E./R^*$  or  $P.E./R'$  (delete the one not used) = 0.0511

hence,  $P.E. = 0.122$  units = 16.3 microbars.  $\therefore L_2 = 52.4 \pm 16.3$  microbars.

Fig. 6.—Section V of the computing form (Section 11(b)).

(iv) Section V is also extended, because the probable error must be determined for each of the four components, lunar semi-diurnal and lunar-solar, namely,  $L_1, \lambda_1$ ;  $L_2, \lambda_2$ ;  $L_3, \lambda_3$ ; and  $L_4, \lambda_4$  (see Section 9).

The whole computation in the present case is too extensive to go conveniently on one sheet. The following alternative suggestions are made for the arrangement of the work.

(1) Use four different sheets, preferably of equal size, each having the necessary particulars\* of Section I entered upon it. The first sheet should correspond to the modified and extended Section II, the second to the extended Section IV, and the third to the extended Section V, as described respectively in (b), (c), and (d). The fourth sheet should include both Section III and Section VI, preferably putting the latter below the former.

\* Full particulars as given in Section I must be indicated on one of these four sheets, though some omissions relating to these particulars may be made for the rest, unless with the aid of carbon papers Section I is completed for all four sheets in one operation.

(2) Furthermore, use the computing form of Figure 3 with slight modification, using subsequently calculated  $\bar{y}$  as the mean of Section II. Use the computing form of the first part of Figure 3,  $\bar{y}_1, \bar{y}_2, \bar{y}_3, \bar{y}_4, \bar{y}_5, \bar{y}_6, \bar{y}_7, \bar{y}_8, \bar{y}_9, \bar{y}_{10}$  as one value on one card each. The 100 computing form sheets (Section II) and the 1000 modification sheets are as follows: Firstly, in Section II the mean  $\bar{y}_1, \bar{y}_2, \bar{y}_3, \bar{y}_4, \bar{y}_5, \bar{y}_6, \bar{y}_7, \bar{y}_8, \bar{y}_9, \bar{y}_{10}$  and  $\bar{y}$  of the differences  $y_1 - y_0, y_2 - y_1, y_3 - y_2, y_4 - y_3, y_5 - y_4, y_6 - y_5, y_7 - y_6, y_8 - y_7, y_9 - y_8, y_{10} - y_9$  and  $y$  are calculated in that order. The subsequently resulting mean subsequent values of  $\bar{y}_1, \bar{y}_2, \bar{y}_3, \bar{y}_4, \bar{y}_5, \bar{y}_6, \bar{y}_7, \bar{y}_8, \bar{y}_9, \bar{y}_{10}$  and  $\bar{y}$  are then calculated. Secondly, the calculation of  $\bar{y}_1, \bar{y}_2, \bar{y}_3, \bar{y}_4, \bar{y}_5, \bar{y}_6, \bar{y}_7, \bar{y}_8, \bar{y}_9, \bar{y}_{10}$  and  $\bar{y}$  from  $\bar{y}_1, \bar{y}_2, \bar{y}_3, \bar{y}_4, \bar{y}_5, \bar{y}_6, \bar{y}_7, \bar{y}_8, \bar{y}_9, \bar{y}_{10}$  and  $\bar{y}$  is made according to Section III, 14) of Section II, and then the 1000th Section of Section II (Figure 1) is also applied. Thirdly, since the value in Section III, 12) and the value  $\bar{y}_1, \bar{y}_2, \bar{y}_3, \bar{y}_4, \bar{y}_5, \bar{y}_6, \bar{y}_7, \bar{y}_8, \bar{y}_9, \bar{y}_{10}$  and  $\bar{y}$  in Section II will have been found for all the data analyzed they are calculated once more, preferably on the last sheet.

*Table 1. Section II of the computing form (Section II).*

1000th Section of Section II			
$\bar{y}_1 = 1.0000$	$\bar{y}_2 = 1.0000$	$\bar{y}_3 = 1.0000$	$\bar{y}_4 = 1.0000$
$\bar{y}_5 = 1.0000$	$\bar{y}_6 = 1.0000$	$\bar{y}_7 = 1.0000$	$\bar{y}_8 = 1.0000$
$\bar{y}_9 = 1.0000$	$\bar{y}_{10} = 1.0000$	$\bar{y} = 1.0000$	$\bar{y} = 1.0000$

Fig. 1. Section II of the computing form (Section II).

## 11. 4) Computing Form for the Transformed (Mean and Variance) Semi-Normal Variations

In this case the computing form as described in Section II is used, with modification for determining both the mean semi-normal mean time and variance (the latter semi-normal computed by the  $F_0$  layer). The latter is described by a slight additional calculation as described in Section 10.

### 11. 4a) Means of Treating the Computations

For each computing sheet such as Figure 1, processing should be as follows:

(i) The sum of the number of fully equivalent in the results, mean of position must be equal to the total number of fully equivalent in the difference in set, i.e.  $\sum_{i=1}^{10} \bar{y}_i = \bar{y}$ , where  $\bar{y}$  denotes the mean, with respect to, from 0 to 1.

(ii) Similarly,  $\sum_{i=1}^{10} \bar{y}_i^2 = \bar{y}^2$ , for every value of  $\bar{y}$  from 0.0000 to 1.0000 when data are used.

(iii) If the collection of set includes all days in each month, month and year, if the initial data include also quantity mean for each hour, the following be used in combination to obtain the sum  $\bar{y}_i = \bar{y}$ .

(iv) Also,  $\sum_{i=1}^{10} \bar{y}_i^2 = \bar{y}^2$ , and  $\sum_{i=1}^{10} \bar{y}_i^3 = \bar{y}^3$ , for each  $\bar{y} = 0.0000, 0.0001, \dots, 1.0000$ .

These sheets of value only in Section II and III. Further important sheets relating to the "load" sheet (Section II) and the separate sheets can be applied as follows:



(v) The numbers  $A_{p, N}$  and  $B_{p, N}$  ( $p=1, 2, 3, 4$ ) in Section III of the "total" sheet must be equal to the *sums* of the corresponding results on its separate set-sheets.

(vi) Similarly, the above rule of checking holds good for such quantities as  $A_{p, A}$ ,  $A_{p, B}$ ,  $B_{p, A}$ ,  $B_{p, B}$ ,  $N_{1, A}$ , and  $N_{1, B}$ , all in Section IV.

## 12. FURTHER GUIDANCE FOR PLANNING AN INVESTIGATION BY THE CHAPMAN-MILLER METHOD

In the preceding sections all important practical details of the Chapman-Miller method and its applications to various geophysical data have been fully described; the following are some other remarks which may usefully be taken into account when planning an investigation of this kind.

### 12. (a) *Length of Data Included*

This depends on many factors, such as the geographical position of the station, the nature of the geophysical element considered, the object of the research, the period or periods of the data actually available for the station, and so on. In the following the phrase "the necessary length of the data" is used for brevity to indicate the shortest period of observation from which the lunar results both for the whole collection of data and the three seasonal sets ( $J, E, D$ ) would be successfully determined, the criterion of success being usually taken to be that the probable error of each determination must exceed one-third of the amplitude of the lunar harmonic considered (Section 10).

Generally speaking the latitude of the station will determine the length of the necessary period, which for meteorological data will be longer the further the station is from the equator. As a matter of fact the discovery of the lunar air-tide originated at tropical stations, where the tide has its largest amplitude, and the other barometric changes are less than in higher latitudes.

The length of the necessary period depends also on the nature of the geophysical element considered. For instance, in the meteorological case it is easier to detect the tidal effect on barometric pressure than that on wind or, still more, on the air temperature. As the effect of the lunar tide in the ionosphere (Section 1) is large, it can be determined from shorter series of ionospheric data.

Further, the length of the necessary period will depend on the nature of the research. For instance, it requires a much longer series of data to investigate the relation of the lunar results to the lunar distance than to the seasons. Clearly the more sets into which we wish to subdivide the data, the greater must be the amount of data used. For instance, at a given station the 12 monthly results derived from a given period of data are always less reliable than the three seasonal ( $J, E, D$ ) results derived from the same length of data; if the latter just reach the standard of acceptance, division into months would not be justified.

Thus as a general rule we should make use of as much data as possible, in order to obtain well-determined results. The more extensive the data included in the investigation, the smaller the probable errors of the derived results, approximately as  $1/\sqrt{N}$ , where  $N$  denotes the number of daily sequences thus used (Section 4).

However, if certain portions of the data appear to be unreliable owing to either instrumental or personal factors, they should be rejected. Similarly, in the meteorological and other cases it may be well to omit those daily observations with too big daily ranges arising from special weather irregularities. All of these, owing to their rather "non-random" characters, will be likely to vitiate the determination of the small periodic variations sought; therefore it is desirable to note and exclude them by a general examination of the data in the planning stage of the investigation.

If the data thus collected still cover a period of 20 or more years, they are usually divided into sub-periods of about 10 years, and treated *independently* to determine lunar results for each sub-period and its three 4-monthly seasons. This practice, when followed, is advantageous, because it enables the results to be compared for different sub-periods, which should show *reasonable agreement*. Then the corresponding results of all sub-periods are combined together, in order to reduce the probable errors involved in each sub-period, giving better-determined corresponding results for the whole period of data thus collected.

#### 12. (b) *Sets Selected According to Other "Character" Figures*

As the intensity of the moon's tidal force upon the earth varies by 20 per cent. with the moon's changing distance, a corresponding variation of the lunar harmonics of all geophysical data may be expected. In order to investigate such dependence, the daily-sequences (Section 4) are usually divided into four nearly equal sets according to lunar distance; these are referred to as PERigee, RECeding, APOgee, and NEARing, corresponding to the four quarters of an anomalistic month, from one perigee to the next, of 27.5546 days. In practice, the quarters PER and APO comprise the seven days centred at the dates of mean perigee and apogee as tabulated in the Bartels and Fauselau's *Mond-Tafeln*, and the quarters REC and NEA are the six or seven days between, on the average, 6.777 days.

Again, the daily sequences may be divided into sets according to lunar declination, in order to study the lunar declination effect upon the lunar reductions thus made. The usual three declination sets are formed according as the mean declinations for corresponding daily-sequences are greater than  $13^{\circ}$  N., between  $13^{\circ}$  N. and  $13^{\circ}$  S., and greater than  $13^{\circ}$  S.

In the cases of geomagnetic and similar data, subdivisions are also made of the whole material into its sets according to the magnetic activity, or the sun-spottedness, or both. These classifications are more or less complicated, depending on the criterion adopted for such "characterization" of daily-sequences, also on the nature and the availability of the actual data involved. For general guidance reference may be made to *Geomagnetism*, Chapters VIII and XI.

#### 12. (c) *Summary of Points and Steps*

The following are some points and steps to be considered or followed in planning an investigation by the Chapman-Miller method:

- (i) Collection of all available data.
- (ii) General examination of the data collected.

- (iii) Exclusions and omissions made for those portions of the data which appear to be unreliable or abnormal.
- (iv) Decision made as to "character" figures to be used for the selection of sets.
- (v) Use of Hollerith cards or tables of daily sequences?
- (vi) Daily sequences formed.

TABLE 6

Section	Symbols used in This Paper	Symbols used by Chapman-Miller(9)
4	$H'; S; f_s; s;$ $\alpha; \beta; \gamma;$ $\mu;$ $C; C'; C'' \dots$ $J; E; D$	$H'; S; f_s$ and $f_{sJ}; s;$ $[\mu_J]$
5	$r; N_r; N;$ $g_s, r; g_s, N$	$r; N_r; N;$ $g_{sr}; g_{sN}$
6	$A_{p', r}; B_{p', r}; p;$ $A_{p', N}; B_{p', N}$	$A_{pSR}; B_{pSR}; p;$ $A_{pSN}; B_{pSN}$
7	$N_{1, A}; N_{1, B}; R;$ $A_{p', A}; A_{p', B}; B_{p', A}; B_{p', B}$	$N_{1RA}; N_{1RB}; R;$ $A_{pSA}; A_{pSB}; B_{pSA}; B_{pSB}$
8	$U_p; V_p; K;$ $L_n; \lambda'_n; \lambda_n;$ $d_{mpS}; m; M; L; L';$ $P; \tau; n; D_{mpS}; t;$ $L; \lambda.$	$X_{pSNR}=U_p/K; Y_{pSNR}=V_p/K;$ $c_m; \lambda_m$ and $\lambda'_m; \lambda_m$ and $\lambda_{mP};$ $d_{mpS}; m; M; L; L';$ $P; \tau; n; D_{mpS}; t$
9	$S_p; \sigma'_p; \sigma_p$	
10	$\Delta a_{p', r}; \Delta b_{p', r}; R^*; L^*; A_r;$ $\Delta a'_{p', r}; \Delta b'_{p', r}; B_r; C_r; D_r;$ $\Delta a''_{p', r}; \Delta b''_{p', r}; R'$	
11	$d; \Delta N_r; \Delta A_{p', r}; \Delta B_{p', r}$	

- (vii) Decision made as to the subdivision of the whole period into its sub-periods.
- (viii) Choice of computing form used.
- (ix) Actual computation for each set on each computing sheet.
- (x) Combination of results for all sub-periods.

## 12. (d) Summary of Characteristics of the Chapman-Miller Method

Some characteristics of the Chapman-Miller method may be summarized as follows :



(ii) It is a fixed age method adapted to determine the short and long solar daily variations in geophysical elements, being series of solar hourly or bi-hourly values covering considerable periods.

(iii) Allowance is made for the "solar 7" or two-apex effect—Section 8 in each day.

(iv) Solar harmonics are derived in the course of the derivation of the lunar harmonics.

(v) The determination of probable errors is easy.

(vi) The use of group-suns reduces the making of numerical computations possible and easy. Equal weight is given to all the observations used.

(vii) Because of its wholesale treatment, it can be applied with machinery unskilled labour working under proper direction. Also the maximum use of Hollerith machines can be made, if the machines are available.

### 12. COMPARISON OF SYMBOLS USED

A comparison of symbols used by the present author and Chapman (1940) is made in Table 6.

### 13. ACKNOWLEDGMENTS

I wish to express my deepest gratitude to Professor S. Chapman, F.R.S., for his supervision and guidance during the course of this work. I am also indebted to the British Council for the provision of a scholarship for post graduate study in Britain, and to the Division of Meteorology, Academia Sinica, China, for granting me leave of absence without such provision and salary support. I would not have been able to carry out this work in this country.

### 14. REFERENCES

- (1) CHAPMAN, S. and TSCHU, K. K. —*Proc. Roy. Soc. A* 195, 324 (1948), where further references are given.
- (2) CHAPMAN, S.—*Ibid.* 137, 1 (1932).
- (3) CHAPMAN, S.—*Memorial Volume, Proc. Fourth Conference of Meteorology, Washington, 1934, 2nd Series, 2nd Meeting, 1935, 1936*.
- (4) APPLETON, E. V., and WILKIN, K. —*Proc. Roy. Soc. A* 171, 171 (1940).
- (5) DE VRIES, A.—*Nature* 137, 296 (1936).
- (6) CHAPMAN, S., and WILKIN, K.—*Compendium of the I. S. U. and I. C. Council* 1940.
- (7) MARTYN, D. F.—*Proc. Roy. Soc. A* 159, 241; *Ibid.* 190, 273 (1947).
- (8) BOUDAL, J., and FRIEDMAN, J. E. —*Geophys. Publ. Ohio*, 11, No. 16 (1947).
- (9) CHAPMAN, S., and WILKIN, K. L. W. —*Rev. Mod. Phys.* 18, 101 (1946).
- (10) BARTLES, J.—*Radio. Geophys.* 34, 36 (1936).
- (11) KENNEDY, I., and KENNEDY, W. T.—*Proc. Roy. Soc. A* 165, 166, 168 (1938).
- (12) CHAPMAN, S.—*Ibid.* 47, 279 (1942).
- (13) MILLER, J. C. P.—*Mém. Nat. R. Acad. Sci. Suppl.* 90, 500 (1934).
- (14) SWEET, J.—*Trans. Amer. Soc. Civil Engrs.* No. 507, 100, 101 (1938).
- (15) BARTLES, J., and FANSELAU, G.—*14th. Geophys. and Petroleum Res. Conf.*, No. 10, 20, 21 (1937). *Trans. Amer. Soc. Civil Engrs.* 101, 102 (1938).

# REGULAR DISTRIBUTION OF PARTICLES FROM THE OF A HOT PLASMA

BY L. M. BAKHAROV, A. L. BAKHAROV, D. D. BAKHAROV and E. BAKHAROV

Page 2, 2.

Summary of report January 15, 1955.

Summary

The regular distribution of particles from the hot plasma is considered. The regular distribution of particles from the hot plasma is considered. The regular distribution of particles from the hot plasma is considered.

where  $kT$  is the temperature of the plasma. The law of the regular distribution of particles from the hot plasma is considered.

## 1. INTRODUCTION

The study of the regular distribution of the particles of a plasma is of interest for the theory of the distribution of particles from the hot plasma. The study of the regular distribution of particles from the hot plasma is of interest for the theory of the distribution of particles from the hot plasma.

The study of the regular distribution of the particles of a plasma is of interest for the theory of the distribution of particles from the hot plasma. The study of the regular distribution of particles from the hot plasma is of interest for the theory of the distribution of particles from the hot plasma.

$$T(T) = T_0 + kT$$

where  $T$  is the temperature of the plasma,  $T_0$  is the temperature of the plasma,  $k$  is the Boltzmann constant,  $T_0$  is the temperature of the plasma,  $k$  is the Boltzmann constant,  $T_0$  is the temperature of the plasma,  $k$  is the Boltzmann constant.

Let us assume that  $T_0$  is the temperature of the plasma,  $k$  is the Boltzmann constant,  $T_0$  is the temperature of the plasma,  $k$  is the Boltzmann constant,  $T_0$  is the temperature of the plasma,  $k$  is the Boltzmann constant.

1.38 MeV. The measurements of Heydenburg *et al.*(7) lie between 1 and 3.5 MeV, but suggest a resonance peak at about 1 MeV. These latter measurements are more precise than those on the low energy side of 1 MeV., and indicate that the angular distribution is given more accurately by

$$Y(\theta)/Y(90^\circ) = 1 + A(E) \cos^2 \theta + B(E) \cos^4 \theta$$

where  $B(E)$  is negative and of the order of one-half between 1 and 3 MeV.

Inglis(2) has pointed out that the presence of the  $\cos^4 \theta$  term implies  $f$  as well as  $p$  protons in the nuclear reaction, and has shown that the width of the peak in the  $A(E)$  curve is determined in part by the extent to which  $f$  protons as well as  $p$  protons are responsible for the reaction. In accounting for the energy dependence of  $A(E)$ , Inglis assumed the presence of two levels in the compound nucleus  $\text{Be}^8$ , one, with angular momentum, 0, possessing a width greater than the range of proton energies involved, the other, with angular momentum, 2, sufficiently narrow to explain the variation of  $A(E)$  near 1 MeV.

The accurate form of the experimental  $A(E)$  curve is thus of theoretical interest and the present paper describes measurements between 100 kV. and 1 MV., made with the aim of improving the accuracy of the data for proton energies below 1 MeV. Nuclear emulsion photographic plates were used to detect emergent  $\alpha$ -particles in the angular ranges 13° to 90°, and 90° to 167°. The values of the ratio  $Y(\theta)/Y(90^\circ)$  were obtained for a somewhat greater range of  $\theta$  than hitherto, a feature of importance in the evaluation of  $B(E)$ . Some 20,000 tracks were counted for each proton energy.

## II. APPARATUS

(i) *Electrostatic Generator.*—Homogeneous beams of protons with energies up to 1 MeV. were produced with the aid of an electrostatic generator and magnetic analyser. The high potential electrode and insulating structure (see Plate I) follow generally the design developed by von Ardenne(8). The electrode is a shell of  $\frac{1}{8}$  in. aluminium and has the form of a cylinder 3 ft. 3 in. long capped with hemispherical ends 5 ft. 3 in. in diameter. This shell is supported internally by an angle-iron framework which is attached to a vertical cylinder of bakelized paper 25 in. in diameter and 0.4 in. in wall thickness. The overall length of the insulating column is 10 ft. 6 in. but it projects through the wall of the shell leaving an insulating length of 9 ft. 2 in. between the surface of the high potential electrode and an earthed steel base. Further support for the high potential electrode, whose axis is horizontal, is provided at one end by two tubular struts mounted at an angle of 30° with the vertical. Originally these struts were bakelized paper tubes 4 in. in diameter and  $\frac{1}{4}$  in. in wall thickness, but as the leakage current along them ultimately became excessive (approximately 50  $\mu\text{A}$ . at 800 kV.) they were replaced with porcelain pipes of similar dimensions.

The bakelized paper insulators are operating in much greater electrical fields than is normal in commercial practice and they have given a good deal of trouble. After twelve months' service the insulation broke down completely. The main insulator was saved by scraping the varnish from the inside and outside surfaces, and after baking it for some weeks under reduced pressure, revarnishing



both surfaces with clear glyptal. At the same time a refrigerating unit was installed in the laboratory, which maintains a humidity of 45 per cent. at  $60^\circ\text{F}$ . At present the leakage current down the main insulator at 800 kV. is approximately 20  $\mu\text{A}$ . over the inside surface and 10  $\mu\text{A}$ . over the outside surface. The leakage along each porcelain strut is 0.25  $\mu\text{A}$ .

The charging belt is of two-ply cotton fabric embedded in pure rubber. The fabric has an open texture which the rubber readily penetrates and yields a finished belt with a continuous layer of rubber about 1 mm. thick on each side of the fabric. The belt is 40 cm. wide and approximately 7.5 m. long.

The charging belt runs on chromium plated steel rollers 10 cm. in diameter. The upper roller is mounted on springs to reduce vibration in the high potential shell. The lower roller, together with the driving motor, is mounted on a platform which can be raised and lowered by four supporting screws which are turned simultaneously by a chain and sprocket drive. This adjustment makes it possible to obtain adequate tension in belts of slightly different lengths. When the belt is travelling at its maximum speed, 1500 m. min., the tension required to prevent slip on the driving roller is approximately 600 lb. weight.

To prevent lateral motion of the belt, the rollers are tapered slightly at the ends and the lower roller is held in a cradle which may be tilted about a horizontal axis. The best alignment of a belt is obtained if it is manufactured with a width in excess of that required and the edges are trimmed subsequently with the belt running.

The belt is driven by a 3 h.p. motor whose speed can be varied continuously to a maximum of 1500 r.p.m. This control is achieved with a Ward-Leonard system of power supply.

With a belt speed of 1500 m. min. and a conventional current-doubler charging system the earthed electrode current is normally about 260  $\mu\text{A}$ .

(d) *Acceleration Tube*.—The acceleration tube is of the cascade type consisting of 12 similar accelerating sections. The general design of the tube will be clear from Plate 1, and the relevant constructional details and dimensions are shown in Figure 1. The final form of the metal part of the tube was determined from field plots made with an electrolytic trough, and is such that the equipotential surfaces are everywhere perpendicular to the glass wall of the tube. In twelve months' service the tube has experienced many violent discharges at the highest potentials but these have passed always between the corona rings. On no occasion has a spark passed to the glass.

The glass sections, *A*, are 23 cm. long, have an external diameter of 25 cm., and an average wall thickness of 5.5 mm. The ends of these glass sections were ground flat and normal to the axis with an accuracy of 0.015 cm., and mild steel flanges, *B* and *C*, were cemented with pinein wax to each end. A small triangular groove, *D*, at the junction of the flanges was filled with Apiezon Q smeared over with vacuum grease to form a vacuum seal between two sections of the tube (see Fig. 1).

The acceleration electrodes, *E*, are of aluminium tube 8 cm. in diameter and each is supported by a light spider machined in the flange, *C*, of each section.

To ensure that the electrodes would remain aligned after the tube was assembled, the end flanges of each section were maintained in axial alignment in a jig during the waxing process, any irregularities of the glass sections being taken up in the piccin wax seal. The alignment of the electrode system was held sufficiently accurate to make unnecessary, at the top of the acceleration tube, devices for aligning the ion source.

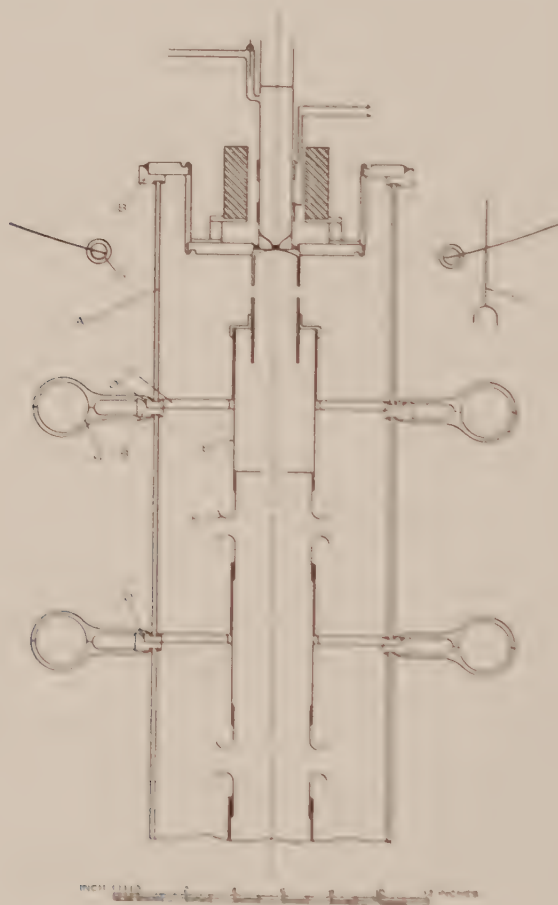


Fig. 1.—Details of accelerator tube.

Aluminium spinnings, *F*, fit tightly into each end of an acceleration electrode and provide a means for adjusting the length of the gap between adjacent electrodes. The ten lower gaps are each 2.5 cm. long, the second 2.0 cm. long, and the first, between the ion source and the first electrode, is 1.1 cm. long.

The acceleration tube was designed so that the plate which carries the ion source is in the plane of the surface of the high potential shell. The edge of the hole through which the tube projects is covered with a rubber gasket, *G*, formed from heavy gauge pressure tubing and has proved to be quite free of corona.

The corona rings,  $H$ , were made in four pieces from aluminium castings which were subsequently machined to shape. When clamped on to the flanges by flush fitting nuts and bolts, the corona rings cover the vacuum seals and clamp the sections rigidly together.

The construction of the target end of the acceleration tube is shown in Plate 2. A steel tube, 19 cm. internal diameter, connects the bottom end of the acceleration tube with a 6 in. diffusion pump (National Research Corporation H-6 Type 104, rated speed 590 l./sec. at  $10^{-4}$  mm. Hg) which is backed with a Hyvac-100, "Narcoid" which is recommended by the makers for use with the diffusion pump proved to have a relatively high vapour pressure and produced a heavy brown deposit in the acceleration tube. The pump functions perfectly with Apiezon B and no further trouble with oil vapour in the tube has been experienced. Measurements made at the bottom of the acceleration tube, after the introduction of a baffle in the pumping line, showed a pumping speed of 140 l./sec. at  $10^{-4}$  mm. Hg. This speed appears to be adequate as it is possible to run the tube steadily at 800 kV., thirty minutes after starting up the pump. Only for operation in the range 900 to 1000 kV. is a short period of "running in" necessary. The pressure at the base of the acceleration tube is observed continuously with an ionization gauge. Normally the tube operates at a pressure of  $2-4 \times 10^{-5}$  mm. Hg.

To fulfil the power requirements in the high potential electrode it was decided to use mechanical methods as far as possible for focusing the ion beam on the target. A measure of focusing is achieved by applying the total potential across an appropriate length of the tube. For example, at 1 MV, the whole length is utilized, while at 100 kV. only the upper two sections are used, the lower ten sections being joined together and connected to earth. Fine control of the focusing is then achieved by adjusting the position of a small cup-shaped conductor,  $J$ , which is connected with the high potential shell and which can be moved vertically between this and the first corona ring (see Plate 1 and Fig. 1). This motion is transferred from a gear box which also controls the ion source and is operated by a long paraffin rod mounted within the main insulator. For potentials ranging from 300 to 1000 kV. the ion beam can be readily focused to produce on a quartz plate the typical brilliant star of light which represents a focal spot about 1 mm. diameter. Even at lower potentials, when the ion beam passes through many feet of field free space before striking the target, focal spots about 2 mm. in diameter are obtained.

For the resolution of the ion beam into its components an electromagnet was used, equivalent in effect to a quadrant of the well-known design of Cockcroft(9), used in the study of x-ray spectra. The magnet is shown in Figures 2 and 3, and in Plate 2. It has the following advantages:

- (i) It is economical in iron and power. With a radius of curvature of 30 cm. a flux density of 4817 gauss is sufficient for 1 MeV. protons. This field is maintained across the 1 cm. gap with the total expenditure of only 220 watts.
- (ii) The resolution is high.



- (iii) The selected homogeneous beam emerges horizontally, which is usually an advantage in experiments.

The emerging end of the electromagnet consists of 2072 turns of 18 S.W.G. enamel-covered copper wire which after being formed to the shape of the core was taped and vacuum impregnated with insulating compound. It is energized by a short wound self-excited D.C. generator which provides a maximum output of 5 A. at 110 V. The generator is driven by a 2-77 h.p. 5-phase induction motor and the output is normally constant, but provision is made for manual fine control by resistors in series with the field windings of the generator.



Fig. 1.—Type and of thermion gun and magnet assembly.  
 G, Thermion gun housing; M, Magnet; Q, Quartz plate; V, Valve;  
 L, Lamp; T, Thermion gun.

The ion beam passes between the pole faces through a quartz plate vacuum box, the sides of which are mild steel plate 8-12 mm. thick and whose top and bottom are copper plates. Gaskets are furnished inside the box (Fig. 1) to prevent outgassing ions emerging from the sources. The vacuum box has flanges over the 2 cm. gap of the magnet, giving an effective gap of 2 cm.

The magnetic analyser is supported on two sets of roller bearings which are carried by a suitable framework and allow two degrees of freedom in the horizontal plane. This permits the ion beam to be centred on the entrance slit of the quarter circle vacuum box.

(iii) *Ion Source*.—The first ion source tried was of the low voltage arc type. Hydrogen gas fed into a cylindrical molybdenum box was ionized by an axial beam of electrons provided by a tungsten filament and accelerated by a potential difference of 500 V. Ionizing efficiency was increased by mounting a cylindrical permanent magnet concentric with this box. The flux density of the longitudinal field in the box was 400 gauss. Positive ions formed inside the box were extracted through a hole in the bottom with a probe potential of some hundreds of volts,

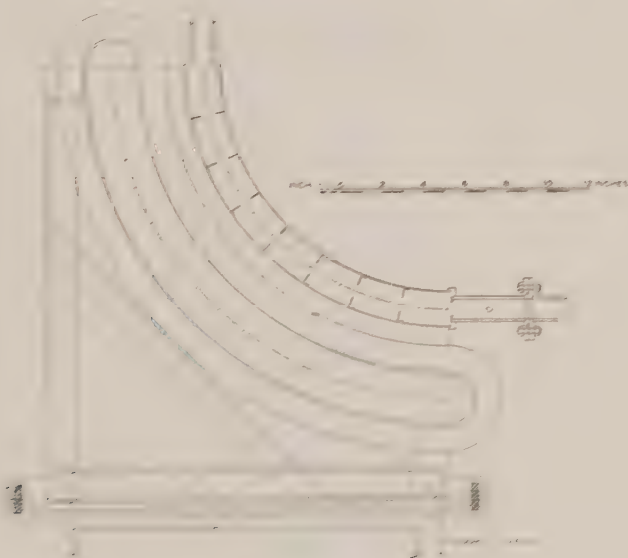


Fig. 3.—Magnetic analyser.

and then passed through a small hole into the accelerating tube. This source gave a total ion current of 12  $\mu\text{A}$ , when operated by batteries placed inside the high potential terminal of the generator. Unfortunately protons represented only some 15 per cent. of the total ion beam.

An Oliphant type source was tried with discharge potentials of 15 kV. and of 30 kV. The energy spread of the protons entering the accelerating tube from this source was considerable and it proved impossible to focus the ion beam satisfactorily. The proportion of protons in the ion beam was approximately 50 per cent.

At present an R.F. type source (10, 11) is being used with satisfactory results (see Fig. 1). This source is energized with a 50 Mc s. 50 W. push-pull oscillator and a probe potential of 3 kV. is used to extract the positive ions. The proton yield is better than 60 per cent. of the total ion current and the energy spread is low. The power necessary to operate the oscillator and to provide the probe

potential is derived from 1000 cycle 80 volt A.C. and 24 volt D.C. generators belt driven from the upper roller.

(iv) *Potential Control and Measurement.*—The potential of the generator was measured with a rotating voltmeter. The design follows one due to Henderson, Goss, and Rose(12) and is shown in Figure 4. The split rotor of the voltmeter is an aluminium cylinder 15 cm. in diameter mounted parallel to the axis of the high voltage electrode and at a distance of 2.7 m. from its surface. The rotor is driven at 1500 r.p.m. by a synchronous motor.

For continuous observation of the potential the A.C. output from the rotor is rectified by a dry rectifier and measured with a moving coil galvanometer shunted to give the required sensitivity and suitable damping. More accurate determination of voltage is made with a 5-dial potentiometer connected across a resistance in series with the rectifier. Potentials can be measured to within 0.5 kV. Calibration of the rotating voltmeter at the standard  $\gamma$ -ray resonances of lithium and fluorine showed that the relation between generator potential and potentiometer reading is linear.

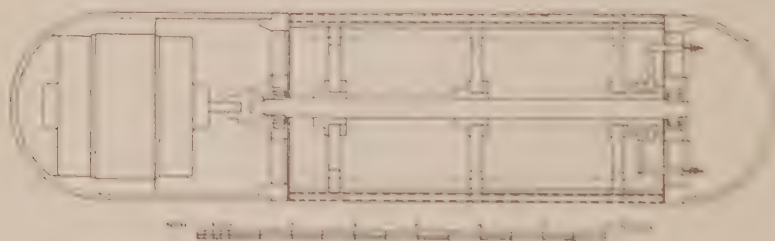


Fig. 4.—Rotating voltmeter.

Control of the potential is effected by a rigid earthed mast pivoted on the floor 3.7 m. distant from the base of the electrostatic generator. Fine control is obtained by changing the angular position of a light metal pointer 25 cm. long mounted at the top of the mast. The metal pointer is rotated by means of a selsyn motor operated at the control table. Between 300 and 800 kV. potentials can be maintained constant to 0.2 per cent.; above 800 kV. to 0.5 per cent.

### III. ANGULAR DISTRIBUTION EXPERIMENT

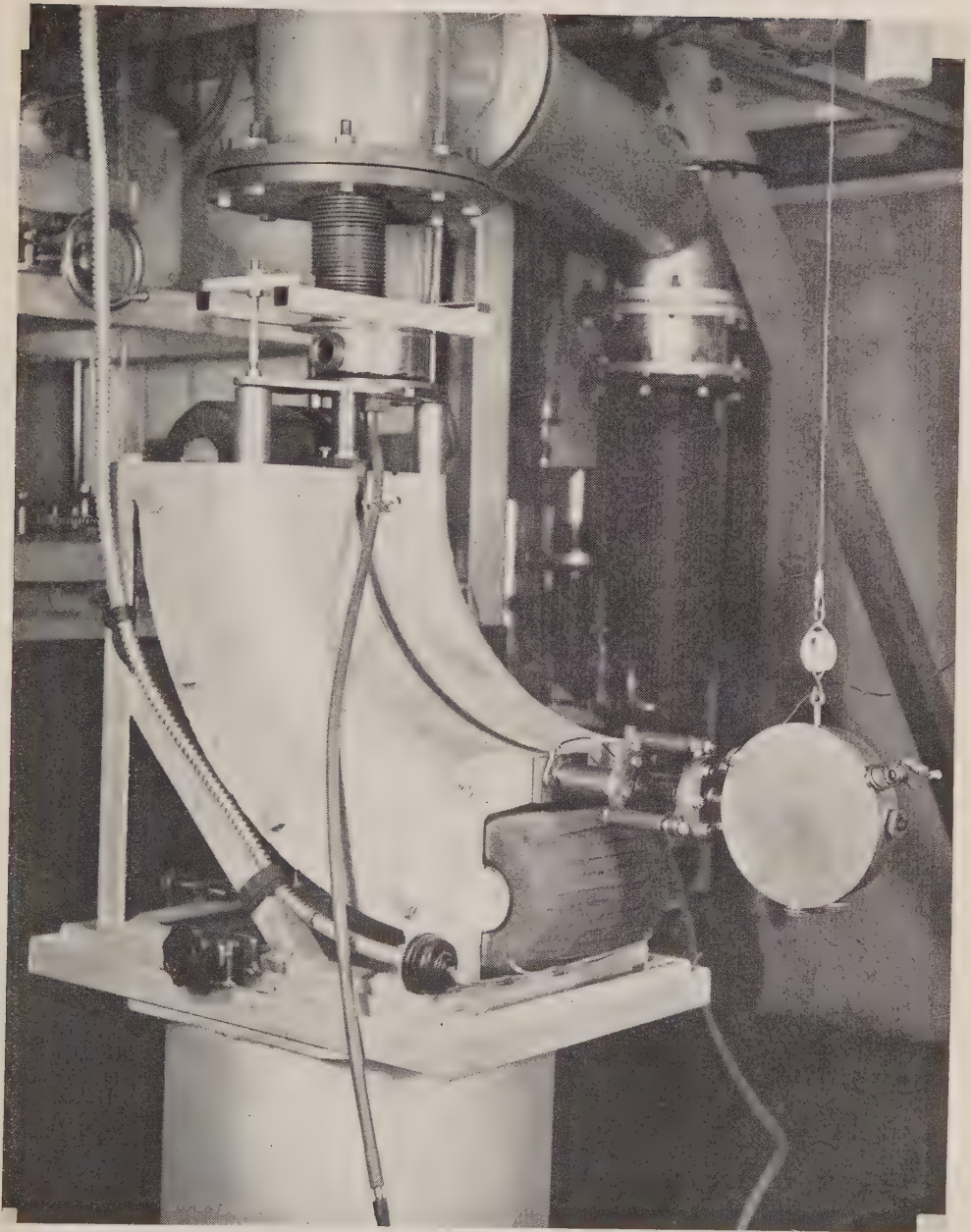
(i) *The Camera.*—The  $\alpha$ -particles from the  $\text{Li}^7(p, \alpha)\text{He}^4$  reaction have been detected using 50  $\mu$  thick Ilford Nuclear Research Emulsion plates (Types C12 and E.1) in the arrangement shown in Figure 5. The incident proton beam enters the target chamber through the tube, *A*, which is tapered so that  $\alpha$ -particles may be observed at angles up to  $167^\circ$  to the incident beam. The beam is collimated by the two circular holes,  $S_1$  and  $S_2$ , whose diameters are 3 and 1.5 mm. respectively. *B* is a small quartz window which facilitates the alignment of the camera. The target was formed by evaporating lithium hydroxide on to aluminium leaf affixed to the aluminium frame, *C*, which was withdrawn from the camera whenever a new target was required. Gamma ray resonance measurements on representative targets showed that the target thickness was





MARTIN ET AL.—ANALYSIS OF GASEOUS MIXTURES BY MEANS OF A GAS CHROMATOGRAPH





MARTIN *et al.*—ANGULAR DISTRIBUTION OF  $\alpha$ -PARTICLES FROM THE  $\text{Li}^7(p,\alpha)\text{He}^4$  REACTION





less than 25 keV. The target spot is an ellipse whose axes are 1.5 and 3 mm. The emitted  $\alpha$ -particles strike the photographic plate, *D*, after emerging through the cellophane window, *E*, which has a stopping power of 2 cm. of air and completely absorbs the scattered protons. A spring, *F*, holds the plate, emulsion side down, against the circular framework, *G*, its position relative to the target being determined from the clearly defined black circle produced on the plate by the scattered protons. The plane of the plate is 1 cm. distant from the axis of the beam. With this target arrangement the yield at all angles, from  $13^\circ$  to  $167^\circ$  to the incident beam, may be obtained with a single exposure. The beam current entering the camera is used to estimate the exposure of the plates.

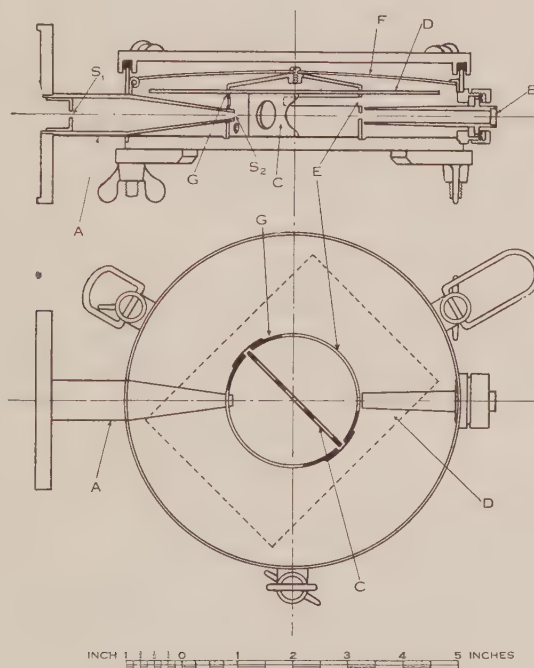


Fig. 5.—Camera.

To prevent excessive heating of the thin target the proton current is restricted to less than  $2\ \mu\text{A}$ . but with the present reaction this current has been quite sufficient.

Although the photographic plate is in the main vacuum system, no difficulty has been experienced either with peeling of the emulsion or in the operation of the accelerating tube.

(ii) *Measurement of the Plate.*—Counts were made along a radial strip and the region of the plate was determined over which the number of  $\alpha$ -particles varied inversely as the square of the distance from the target.

Two concentric circles of radii 4 and 4.5 cm. are scribed on each plate with a fine needle point, and define an area for counting which lies well within

the "inverse square region". Radial lines are scribed on each plate at each angle where the yield is to be measured. In the plane of the plate these angles are  $0^\circ$ ,  $30^\circ$ ,  $45^\circ$ ,  $60^\circ$ ,  $90^\circ$  (on each side of the beam),  $120^\circ$ ,  $135^\circ$ ,  $150^\circ$ , and  $180^\circ$  to the direction of the incident beam.

A binocular microscope, with a  $\frac{1}{8}$  in. objective and  $\times 6$  eyepieces, is used to count the tracks in a suitable area at the selected positions on the plate. The area selected depends on the track density but the angle subtended by it at the target never exceeds  $2^\circ$ . An eyepiece graticule in the form of a square enables an accurate count of the number of particle tracks in a known area to be made. With this arrangement an experienced observer can count 6000 tracks per day without strain. Check counts by independent observers show an agreement of better than 1 per cent. when correction is made for the small change in magnification produced by a change in the interocular distance of the microscope eyepieces. One of the chief advantages of the photographic plate technique is that only those particles from the desired reaction, i.e. those with the appropriate length and direction, are counted.

At least two plates have been exposed and approximately 20,000 tracks counted at each accelerating voltage.

#### IV. RESULTS

After calculating the true angles between the tracks and the incident beam from the angles measured in the plane of the plate, the observations are converted(13) to centre of mass coordinates. Simple geometry shows that the first correction is given by

$$\sin^2 \varphi = \frac{\sin^2 \psi + h^2/r^2}{1 + h^2/r^2}$$

where  $\varphi$  = true angle with the incident beam

$\psi$  = observed angle in the plane of the plate

$h$  = normal distance of axis of proton beam from the plate

$r$  = mean radius of circles scribed on the plate.

These angles in laboratory coordinates are then converted to centre of mass coordinates using the relations

$$\cot \varphi = \cot \theta + k \operatorname{cosec} \theta, \quad k = r/V_2 = \left[ \frac{M_1 M_2 E_1}{M M_3 Q + M_0 M_3 E_1} \right]^{1/2}$$

where  $\varphi$  = angle in laboratory coordinates

$\theta$  = angle in centre of mass coordinates

$v_c$  = velocity of centre of mass in laboratory coordinates

$V_2$  = velocity of  $\alpha$ -particles in centre of mass coordinates

$M_0$  = mass of stationary target nucleus

$M_1$  = mass of incident nucleus (proton)

$M_2$  = mass of observed nucleus ( $\alpha$ -particle)

$M_3$  = mass of residual nucleus

$E_1$  = energy of incident proton in laboratory coordinates

$Q$  = reaction energy

$M = M_0 + M_1 = M_2 + M_3$ .



For convenient calculation, however, it is desirable to express  $\theta$  in terms of  $\varphi$

$$\cos \theta = -k \sin^2 \varphi + (1 - k^2 \sin^2 \varphi)^{\frac{1}{2}} \cos \varphi.$$

Finally, the solid angle subtended by the area, in which tracks are counted, at the target must be converted to the centre of mass coordinate system.

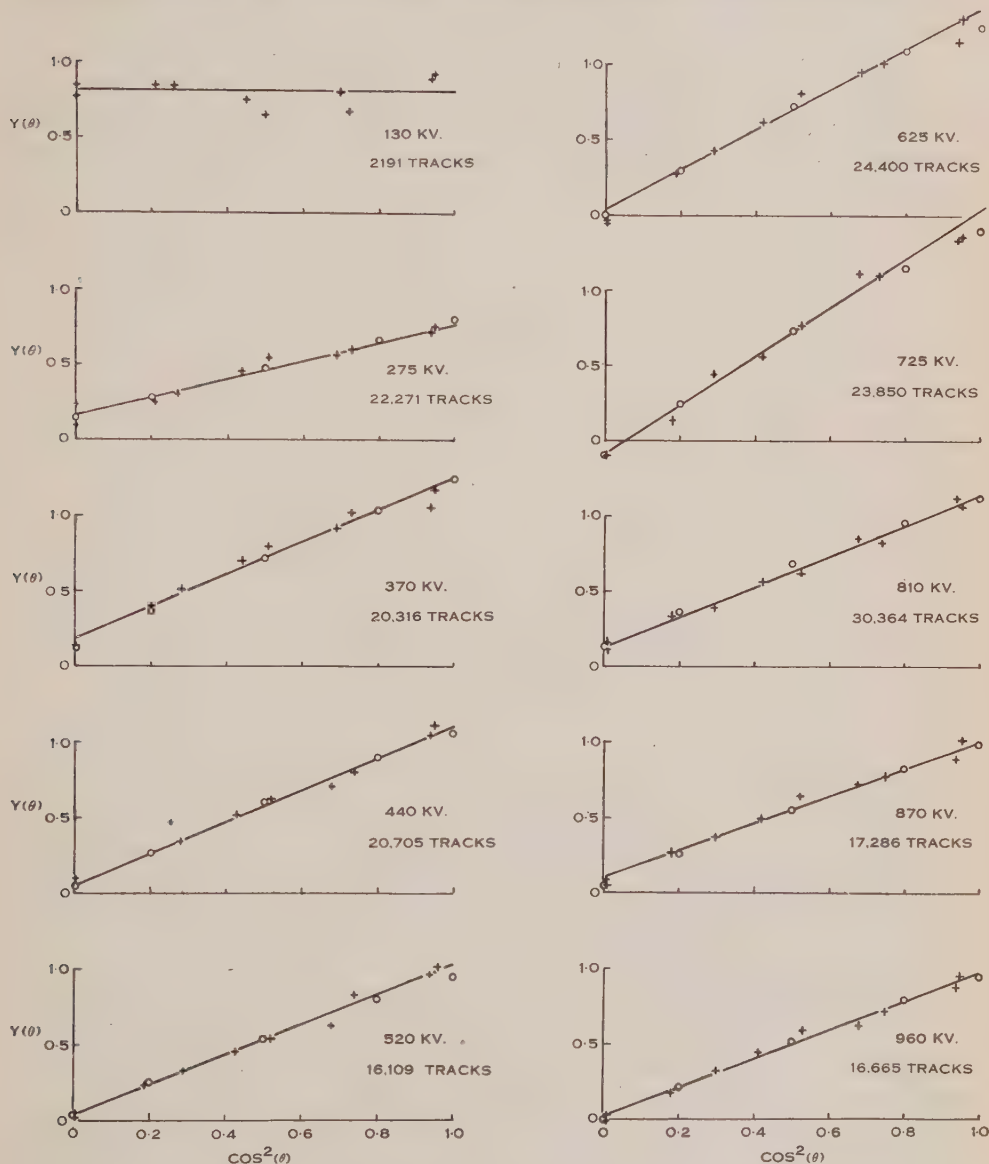


Fig. 6.— $Y(\theta)$  against  $\cos^2 \theta$ .

The corrected yield is given by

$$Y(\theta) = \frac{Y(\varphi)}{2k \cos \varphi + (1 - k^2 \sin^2 \varphi)^{\frac{1}{2}} + k^2 \cos^2 \varphi (1 - k^2 \sin^2 \varphi)^{-\frac{1}{2}}}$$

where  $Y(\theta)$  = yield of  $\alpha$  particles per unit solid angle in centre of mass coordinates

$Y(\varphi)$  = yield of  $\alpha$  particles per unit solid angle in laboratory coordinates.

Figure 6 shows relative values of  $Y(\theta)$  plotted against  $\cos^2 \theta$  for ten different proton energies within the range 100 keV. to 1 MeV. The values of  $A_0(E)$  obtained from the slopes of the straight lines of best fit are shown plotted against  $E$  in Figure 7, together with values of  $A_0(E)$  obtained by previous workers in this range. It will be noticed that our values indicate a maximum for  $A_0(E)$  at energies greater than 950 keV. and are consistent with the results of Heydenburg *et al.* (7) in placing the maximum at about 1 MeV. This energy is rather greater than the value, 675 keV., found by Swartz *et al.* (4), and 850 keV. found by Rubin *et al.* (5, 6).

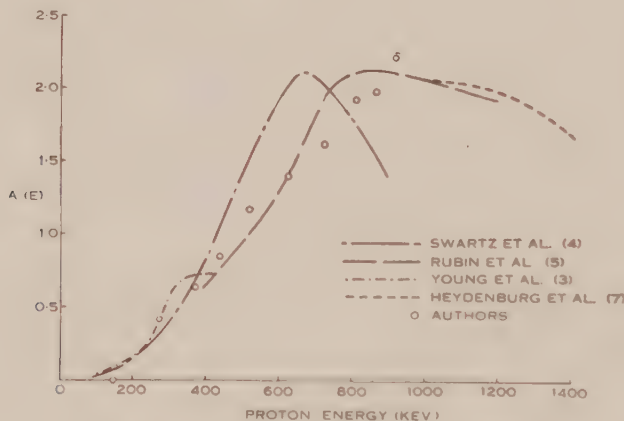


Fig. 7.— $A_0(E)$  against proton energy.

In earlier experiments (3, 4, 5, 6) the apparatus precluded measurements of  $Y(\theta)$  in directions adjacent to the axis of the proton beam with the result that  $\theta$  was limited to values for which  $\cos^2 \theta \geq 0.77$ . The values of  $A_0(E)$  due to Heydenburg *et al.* shown in Figure 7 refer to this range of  $\cos^2 \theta$ . For greater values of  $\cos^2 \theta$  Heydenburg *et al.* found that  $Y(\theta)$  lay below the straight line defining  $A_0(E)$ , and that a further term,  $B(E) \cos^4 \theta$ , was required to describe the asymmetry adequately.

In an attempt to detect the presence of this term for proton energies less than 1 MeV. the camera was designed to record  $\alpha$ -particles at angles for which  $\cos^2 \theta \leq 0.95$ . Although, as was mentioned above, the  $Y(\theta)$ ,  $E$  curve can be represented by a straight line, it will be seen from Figure 6 that there is a persistent tendency for the  $Y(\theta)$ ,  $E$  curves to bend towards the  $\cos^2 \theta$  axis. It is difficult to evaluate  $B(E)$  with any accuracy since it represents a residual of the  $Y(\theta)$ ,  $Y(90^\circ)$ ,  $\cos^2 \theta$  curve.

Values of  $A(E)$  and  $B(E)$  were obtained by fitting the expression

$$Y(\theta) = Y(90^\circ)[1 + A(E) \cos^2 \theta + B(E) \cos^4 \theta]$$

to the experimental points by the method of least squares.

$A(E)$  and  $B(E)$  are shown plotted against  $E$  in Figure 8,\* together with values given by Heydenburg *et al.* in their Table 1 (first paper) and Figure 6 (second paper). We have shown the separate determinations given in their table since

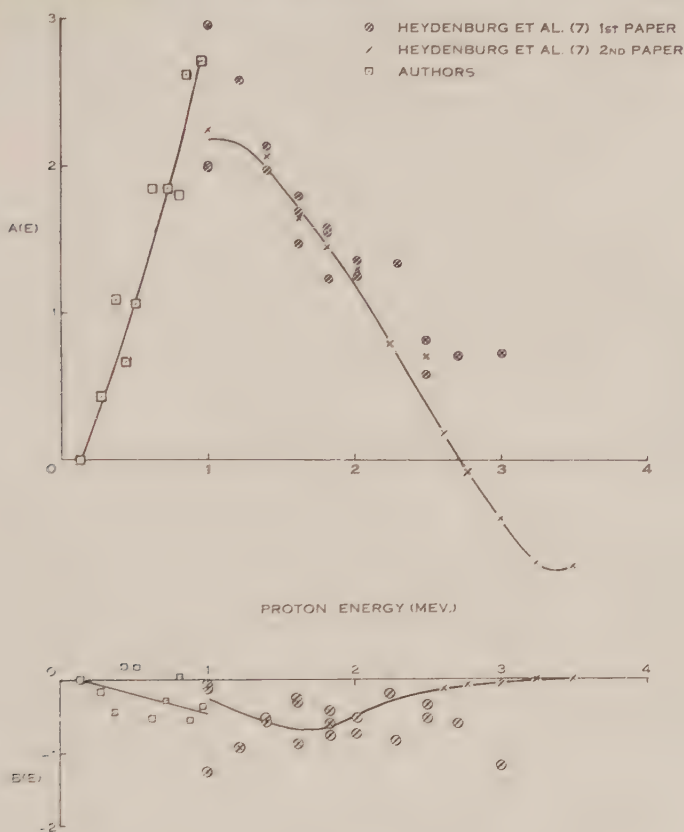


Fig. 8.— $A(E)$  and  $B(E)$  against proton energy.

\* Further examination of the data showed that the magnitude of  $B(E)$  was very sensitive to target thickness. A check on the thickness of each target was made by determining the yield of tracks in a given direction on each plate per microcoulomb of protons, and this showed that some of the targets were thicker than had been estimated by visual comparison with others whose thickness was determined from  $\gamma$ -ray resonances ( $< 25$  keV.). Measurements on the best of the original plates have been extended and have been combined with measurements on further plates exposed using thin targets ( $< 10$  keV. thick).

The final values of  $A(E)$  and  $B(E)$  obtained are :

$E$ (keV.) ..	130	275	370	440	520	625	715	810	870	960
$A(E)$ ..	0.00	0.67	1.00	0.98	1.18	1.55	1.95	2.25	2.68	2.80
$B(E)$ ..	0.0	-0.2	-0.3	-0.2	0.0	-0.2	-0.5	-0.5	-0.6	-0.5

These values lie closely on the smooth curves shown in Figure 8.



our results near 1 MeV. are in agreement with some of their higher values. Heydenburg *et al.* mention that there is no reason known to them for discounting these high values.

The points on the curve in Figure 6 represented by circles have been calculated from the expression  $1 + A(E) \cos^2 \theta + B(E) \cos^4 \theta$  using values of  $A(E)$  and  $B(E)$  taken from the lines drawn in Figure 8. In spite of the uncertainty in the values of  $B(E)$  these points give a better fit to the experimental data than the straight lines which represent  $1 + A_0(E) \cos^2 \theta$ .

We regard the evidence for the existence of the  $B(E) \cos^4 \theta$  term as conclusive, confirming the assumption by Inglis(2) of the presence of  $f$  protons in the lithium reaction at proton energies less than 1 MeV. However, the data at present lack sufficient accuracy to provide precise values for the adjustable constants which appear in the theoretical expressions for  $A(E)$  and  $B(E)$ . While the number of counts at 130 kV. is small, it will be clear from Figure 6 that the value of  $A(E)$  is close to zero, a result also to be expected from theory (2).

## V. ACKNOWLEDGMENTS

The cost of the electrostatic generator was met from a grant made by the Council for Scientific and Industrial Research as part of their research programme in nuclear physics. The authors wish to express their thanks to Dunlop Rubber Australia Ltd. for their help in producing satisfactory charging belts, and for financial assistance. Much of the detail of the electrical and mechanical work was due to Mr. W. G. Caldwell, and the successful operation of the generator is due in large measure to his assistance. The photographic plate measurements have been carried out by Miss G. B. Ward with the assistance of Miss E. V. Sherriff.

## VI. REFERENCES

- (1) CRITCHFIELD, C. L., and TELLER, E.—*Phys. Rev.* **60** : 10 (1941).
- (2) INGLIS, D. R.—*Ibid.* **74** : 21 (1948).
- (3) YOUNG, V. J., ELLETT, A., and PLAIN, G. J.—*Ibid.* **58** : 498 (1940).
- (4) SWARTZ, C. D., ROSSI, H. H., JENNINGS, B., and INGLIS, D. R.—*Ibid.* **65** : 80 (1944).
- (5) RUBIN, S., FOWLER, W. A., and LAURITSEN, C. C.—*Ibid.* **71** : 212 (1947).
- (6) RUBIN, S.—*Ibid.* **72** : 1176 (1947).
- (7) HEYDENBURG, N. P., HUDSON, C. M., INGLIS, D. R., and WHITEHEAD, W. D.—*Ibid.* **73** : 241 (1948); *Ibid.* **74** : 405 (1948).
- (8) ARDENNE, M. VON.—*Z. Phys.* **121** : 236 (1943).
- (9) COCKCROFT, J. D.—*J. Sci. Instrum.* **10** : 71 (1933).
- (10) THONEMANN, P. C.—*Nature* **158** : 1 (1946).
- (11) RUTHERGLEN, J. G., and COLE, J. F.—*Ibid.* **160** : 545 (1947).
- (12) HENDERSON, J. E., GOSS, W. H., and ROSE, J. E.—*Rev. Sci. Instrum.* **6** : 63 (1935).
- (13) MANDEVILLE, C. E.—*J. Franklin Inst.* **244** : 385 (1947).

## EXPLANATION OF PLATES 1, 2

### PLATE 1

Electrostatic generator.

### PLATE 2

Magnetic analyser and camera.

# HALL CURRENTS AND THE EJECTION OF PROMINENCES BY SUNSPOTS

By R. G. GIOVANELLI\*

[*Manuscript received September 24, 1948*]

## *Summary*

During the growth of sunspots induced electric fields may be expected to be set up in the surrounding atmosphere. It is shown that, because of the comparatively low conductivity perpendicular to lines of magnetic force, there are localized regions where large space charges occur, resulting in large electric fields perpendicular to the lines of magnetic force. Consequently both positive and negative charges drift in the same sense in a direction which is at right angles to the electric and magnetic fields, giving rise to a general movement of the gas. The drift velocities are difficult to estimate, but appear to be of the order of magnitude of those found in eruptive prominences.

## I. INTRODUCTION

It is well known that eruptive prominences often occur in the neighbourhood of a chromospheric flare. We present here a theory to explain this phenomenon ; according to this theory the ejection is due to the Hall effect, whereby ions of both signs drift in a direction perpendicular to crossed electric and magnetic fields. We first briefly recapitulate the theory of the origin of these fields and the resulting current system, which was discussed in detail in an earlier paper(1).

## II. ELECTROMAGNETIC FIELDS ASSOCIATED WITH A GROWING SPOT GROUP

It is well established that sunspots normally grow in bipolar groups, the magnetic axis being approximately east-west. We will therefore consider a spot group situated in the normal sunspot zone, and consisting of two sunspots of the same pole strengths but opposite polarities, the axis of the group being perpendicular to the sun's general magnetic field which will be inclined to the sun's surface.

The origin of a group of sunspots is still uncertain. For convenience we adopt a model in which a magnetic toroid, originally below the sun's surface, breaks through the surface to form two regions (the sunspots) of high magnetic field strength and opposite polarity. Results differing in detail but in general similar to those to be described in this paper would, however, be obtained with any other sunspot model.

With this model, there is established during the growth of the group an electric field which, if the sun's atmosphere were a non-conductor, would be directed around circles perpendicular to and concentric with the line joining the

\* Division of Physics, C.S.I.R.

two sunspots. For a sunspot growing in fifty hours to a diameter one-hundredth that of the sun, and having at its maximum size a magnetic field of 2000 gauss which we may conveniently assume to be uniform across the spot, the rate of increase of flux from the spot being assumed constant, the induced e.m.f. around any circuit cut by all the escaping flux is  $1.7 \times 10^{16}$  e.m.u. The fraction of the flux which cuts a circular circuit concentric with the axis of the group decreases as the radius increases, so that the e.m.f. is lower for circuits of larger radii. It has been shown that the induced electric field strength at a distance of  $3 \times 10^9$  cm. is of the order of  $10^5$  e.m.u.

The sun's atmosphere is, however, far from being a non-conducting medium. In the absence of a magnetic field, its conductivity would be of the order of  $10^{-8}$  e.m.u. A magnetic field reduces the conductivity at right angles to the lines of magnetic force to such an extent that near sunspots electric currents, except in special cases, must flow along lines of magnetic force. Consequently

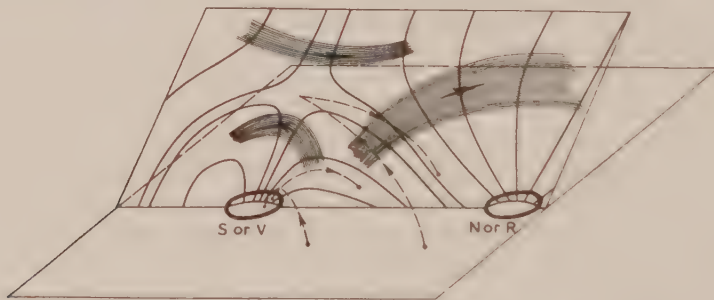


Fig. 1.—Streams near a simple bipolar spot group.

Lines of magnetic force in the dividing plane are shown by continuous lines. Typical current paths are indicated by broken lines, and the direction of the current indicated by arrows for the case where the group is growing with south, or *V*, polarity in the left-hand spot. Possible streams, which must lie in the dividing plane, are shown by the broad dark areas, and the directions of motion indicated.

tubes of magnetic force may, to a first approximation, be regarded as electrical conductors. Circumstances in which this approximation is inadequate will be examined later in this paper.

Accompanying the growth of a sunspot, there will be a system of induced currents in the sun's atmosphere in the neighbourhood of the sunspot system. For the model which we have adopted, Figure 1 shows such a system when the total magnetic flux from the sunspots is not too large, so that all the lines of force from one sunspot are absorbed into the general field and none of them reaches the other spot. Part of the current passes near a neutral point in the magnetic field. Another part of the current flows into a sunspot on one side of the "dividing" plane (the plane which contains the general field and the two sunspots) and flows out from the same sunspot on the other side. This difference in direction of the current on opposite sides of the dividing plane plays an important part in our discussion, and is illustrated in Figure 2. It arises from the way in which the components of the induced electric field along the lines of magnetic force reverse direction on crossing the dividing plane.



We now consider the electric field strength normal to the dividing plane. This depends in part on the distribution of resistance around current paths which lead into or out of sunspots. While the body of the sun may normally be regarded as a good conductor, the conducting properties of a sunspot below the sun's surface are very uncertain, owing to our lack of knowledge of the strength and extent of the intense magnetic field. It is not unlikely that the greater part of the resistance of any complete current filament which passes through the sun's atmosphere into the sunspot lies within the sunspot. In any case, the cross-sectional area of any current filament narrows rapidly on approaching the sunspot, and most of the resistance of a current filament in the atmosphere is near the sunspot. It follows that there is a large potential difference between points on opposite sides of the dividing plane, and the electric intensity across this plane, at right angles to the magnetic field, is much greater than if the medium were non-conducting.

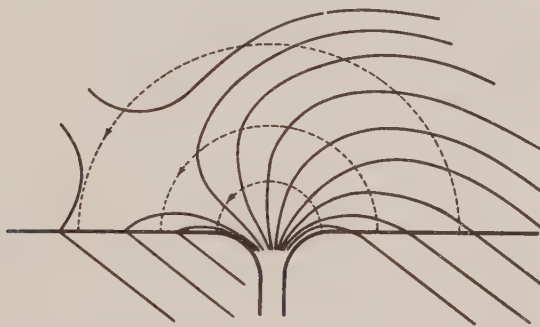


Fig. 2.—Section through sunspot and neutral point. The projections of lines of magnetic force are shown in a vertical plane intersecting a sunspot and a neutral point. The direction of the induced electric field is indicated by broken lines.

The above combination of electric and magnetic fields will give rise to Hall currents, or the movement of charged particles perpendicular to both electric and magnetic fields. Charges of both signs move in the same direction and, in most cases, with equal velocities. The velocity depends, among other things, upon the electric field strength, which in turn depends to a large extent on the value of the small conductivity at right angles to the magnetic field. As is shown below, the velocities in localized regions are expected to be of the order of 100 km./sec. or more.

We may note that in our simple sunspot group the streams of charged particles resulting from this cause will be confined to the neighbourhood of, and move parallel to, the dividing plane.

### III. THE MOTION OF THE IONS IN ELECTRIC AND MAGNETIC FIELDS

It is well known that if in any uniform continuous medium an electric field  $E$  acts at right angles to a magnetic field  $H$ , the drift velocity of charged particles is not necessarily parallel to the electric field, but is in a plane perpendicular to

the magnetic field. The drift velocity may, in fact, be regarded as having two components, one (usually termed the direct velocity,  $w$ ) in the direction of the electric field, and the other (termed the transverse velocity,  $v$ ) at right angles to both the electric and magnetic fields. For non-relativistic cases(2),

$$w = E\omega\tau/H(1 + \omega^2\tau^2)$$

and

$$v = E\omega^2\tau^2/H(1 + \omega^2\tau^2)$$

where  $\omega = eH/m$  ( $e$  and  $m$  being the charge of the particle in e.m.u. and its mass in grams, respectively) and  $\tau$  is the mean time between collisions of the particle with particles of other types. For a highly ionized gas, such as the sun's atmosphere,  $\tau$  is given by(3)

$$\tau = u^3/\pi\beta^2N_+Z^2 \text{ sec.},$$

where  $u$  is the mean electron velocity,  $\beta$  is a constant equal to  $1.57 \times 10^9 \text{ cm.}^3/\text{sec.}^2$ ,  $N_+$  is the number of positive ions per cc., and  $Z$  is their mean degree of ionization (unity in the case of hydrogen).

If  $H$  is very small, then the direct velocity  $w$  equals  $E\tau/m$  and the transverse velocity  $v$  is negligible. For values of  $H$  such that  $\omega\tau \gg 1$ ,  $v$  approaches  $E/H$ , so that for  $H$  large enough the transverse velocity is independent of the nature of the ion.

The ratio  $v/w$  is always given by

$$v/w = \omega\tau = eH\tau/m,$$

so that when  $\omega\tau$  is greater than unity, the transverse velocity exceeds the direct velocity. For example, if the electron temperature,  $\theta$ , is  $3 \times 10^4 \text{ K.}$ , so that the electron velocity is  $1.2 \times 10^8 \text{ cm./sec.}$ , if  $N$  is  $10^{10}$  per cc., and if  $H$  is 100 gauss, then  $v/w$  is  $3.1 \times 10^4$  for electrons and  $1.7 \times 10^3$  for protons.

When  $E = 10^5 \text{ e.m.u.}$  (a typical value of the induced field expected in a non-conducting atmosphere) and  $H = 100 \text{ gauss}$ , the transverse velocity is

$$v = E/H = 10^3 \text{ cm./sec.}$$

Thus velocities comparable with those observed in prominences (say  $5 \times 10^6 \text{ cm./sec.}$ ) occur only where  $E$  is considerably larger, or  $H$  considerably smaller, than these values.

#### IV. THE ELECTRICAL CONDUCTIVITY OF THE CHROMOSPHERE AND CORONA

In order to estimate the current resulting from the electric field, we may resolve the latter along and perpendicular to the magnetic field. The components of the current may be approximately obtained from the corresponding conductivities. Denote the conductivity of the medium in the direction of the magnetic field by  $\sigma_0$ , the direct conductivity at right angles to the magnetic field and in the plane containing the electric and magnetic fields by  $\sigma^1$ , and the transverse conductivity by  $\sigma^u$ . Then if  $n$  be the electron concentration,

$$\sigma_0 = ne^2\tau/m = e^2u^3/\pi\beta^2m,$$

which has the value  $5.0 \times 10^{-8} \text{ e.m.u.}$  if  $u$  corresponds to  $3 \times 10^4 \text{ }^\circ\text{K.}$ , as in the chromosphere, but may be closer to  $5 \times 10^{-5} \text{ e.m.u.}$  in the corona, where the

temperature is higher by a factor of about  $10^2$ . Again, if  $N_+$  be the ion concentration,

$$\sigma^I = \frac{ne\omega\tau}{H(1+\omega^2\tau^2)} + \frac{N_+e\Omega\tau}{H(1+\Omega^2\tau^2)},$$

when the first and second terms represent contributions from electrons and protons respectively, the symbols for protons being similar to those for electrons.  $\sigma^I$  is given for a temperature of  $3 \times 10^4$  °K. and for several values of  $H$  in Table 1. Again,

$$\sigma^{II} = \frac{ne\omega^2\tau^2}{H(1+\Omega^2\tau^2)} - \frac{N_+e\Omega^2\tau^2}{H(1+\Omega^2\tau^2)}.$$

When  $H$  is small, so that  $\Omega^2\tau^2 \ll 1$ , the proton contribution is negligible, but when  $\Omega^2\tau^2 \gg 1$ , both terms become numerically equal and  $\sigma^{II}$  vanishes, provided  $n = N_+$ , as is true to a high degree of approximation. This might also be seen from the equality of the transverse velocities of electrons and protons under these conditions.

TABLE 1

$a/a_n$	$H$ (gauss)	$\sigma^I$ (e.m.u.)	$E_m$ (e.m.u.)	$E_m/H$ (cm./sec.)	$y_m$ (cm.)
0.2	600	$2.6 \times 10^{-15}$	$2.5 \times 10^{10}$	$4.1 \times 10^7$	$4.5 \times 10^5$
0.3	250	$1.5 \times 10^{-14}$	$1.0 \times 10^{10}$	$4.1 \times 10^7$	$1.1 \times 10^6$
0.4	130	$5.5 \times 10^{-14}$	$5.3 \times 10^9$	$4.1 \times 10^7$	$2.1 \times 10^6$
0.5	75	$1.6 \times 10^{-13}$	$3.1 \times 10^9$	$4.1 \times 10^7$	$3.6 \times 10^6$
0.6	45	$4.5 \times 10^{-13}$	$1.9 \times 10^9$	$4.2 \times 10^7$	$6.0 \times 10^6$
0.7	26	$1.3 \times 10^{-12}$	$1.1 \times 10^9$	$4.2 \times 10^7$	$1.0 \times 10^7$
0.8	13	$4.6 \times 10^{-12}$	$5.9 \times 10^8$	$4.5 \times 10^7$	$1.9 \times 10^7$
0.9	5.7	$1.4 \times 10^{-11}$	$3.4 \times 10^8$	$6.0 \times 10^7$	$3.3 \times 10^7$
0.95	2.7	$2.2 \times 10^{-11}$	$2.7 \times 10^8$	$1.0 \times 10^8$	$4.2 \times 10^7$
0.97	1.6	$2.5 \times 10^{-11}$	$2.5 \times 10^8$	$1.6 \times 10^8$	$4.5 \times 10^7$
0.99	0.5	$2.9 \times 10^{-11}$	$2.3 \times 10^8$	$4.6 \times 10^8$	$4.8 \times 10^7$

## V. THE ELECTRIC FIELD NEAR THE DIVIDING PLANE

It is clear that along a line normal to the dividing plane, the maximum electric field strength perpendicular to the magnetic field will occur at the dividing plane. At the neutral point the magnetic field is zero; so that we may expect, and will show later, that the transverse velocity  $v$ , ( $=E/H$ ), has a maximum value in this plane somewhere near the neutral point. The direction of  $v$  lies in the dividing plane. We noted earlier that there is a large potential difference between points on either side of the dividing plane. If, in fact, there were zero conductivity across the lines of magnetic force, there would be a sudden change of potential on crossing the dividing plane, this change being less than, but of the same order of magnitude as, the total induced e.m.f.,  $V_0$ , around possible current paths. However, the actual electric field strength at right angles to the dividing plane will depend on the magnitude of the small but finite conductivity perpendicular to the magnetic field.



It is not feasible to take account of all factors which materially affect the electric field strength. The order of magnitude of the field may be found by comparison with the two-dimensional problem of the electric field when two halves of an infinite plane, separated by a straight line, are maintained at a potential difference  $V_0$ , the conductivity of the surrounding medium in the  $x$ -direction, perpendicular to the plane, being  $\sigma_0$ , and the conductivity in the  $y$ -direction, parallel to the plane and perpendicular to the line of separation, being  $\sigma^1$ .

In such a case the potential at any point is given by

$$\frac{\partial}{\partial x} \left( \sigma_0 \frac{\partial V}{\partial x} \right) + \frac{\partial}{\partial y} \left( \sigma^1 \frac{\partial V}{\partial y} \right) = 0.$$

If  $\sigma_0$  and  $\sigma^1$  are constants, we may transfer to coordinates  $X = x$ ,  $Y = y (\sigma_0 / \sigma^1)^{1/2}$ , whence

$$\frac{\partial^2 V}{\partial X^2} + \frac{\partial^2 V}{\partial Y^2} = 0,$$

the solution of which for the above boundary conditions is well known(4), viz.,

$$V = (V_0 / \pi) \tan^{-1}(X/Y) = (V_0 / \pi) \tan^{-1}\{(x/y)(\sigma^1 / \sigma_0)^{1/2}\},$$

while the electric field parallel to the  $y$ -axis is

$$E = dV/dy = -(V_0 x / \pi) (\sigma_0 / \sigma^1)^{1/2} / (x^2 + y^2 \sigma_0 / \sigma^1).$$

For a given value of  $x$ , the field has a maximum value  $E_m = (V_0 / \pi x) (\sigma_0 / \sigma^1)^{1/2}$  at  $y = 0$ , while the range  $y_m$  of  $y$ , over which the field exceeds half of the maximum, is  $2x(\sigma^1 / \sigma_0)^{1/2}$ .

In our actual problem,  $H$  and therefore  $\sigma^1$  vary considerably throughout the region under consideration: at a distance  $a$  from the sunspot along the line joining an isolated sunspot to the neutral point

$$H = G(a_n^2 - a^2) / a^2$$

where  $G$  is the strength of the general magnetic field and  $a_n$  the distance between the sunspot and neutral point. The values of  $H$  and  $\sigma^1$  for  $G = 25$  gauss,  $\theta = 3 \times 10^4$  K., and  $N = 10^{10}$  per cc. are given for various values of  $a/a_n$  in Table 1. Putting  $V_0 = 1.7 \times 10^{16}$  e.m.u. (a typical value) and  $x \approx 10^9$  cm., the orders of magnitude of  $E_m$  and  $y_m$  may be obtained for the different values of  $\sigma^1$  (see Table 1).

We may note that  $E_m H$  (which gives the velocity of the ions) is of the order of  $5 \times 10^7$  cm./sec. and so is rather larger than, but of the same order of magnitude as, the velocities in eruptive prominences.

## VI. LIMITATIONS ON MOVEMENTS OF IONS

So far we have considered the motion of particles as if their movements had no effects upon the phenomena causing them. It is possible, however, that polarization may enforce limitations upon the whole process, and we now investigate this possibility. In a medium of uniform particle concentration, positive and negative particles can move in a complete circuit with different drift velocities without setting up any local space charge. But in the sun's atmosphere the concentration of particles decreases with height, and in general the velocities of the particles will have some vertical component. If this com-

ponent differed for electrons and protons, the space charge so formed would instantly set up a polarization field tending to prevent further separation. In part, the space charge producing this polarization would be dispersed along lines of magnetic force, but this dispersal will never be complete. The case most unfavourable to mass motion would be that in which the dispersal of charges is negligible; the polarization field would then be such that electrons and protons would move with identical vertical components; this implies that the velocities would themselves be identical.

The polarization field gives rise to a transverse, or Hall, current which is in the same direction as the direct current due to the original field  $E$ . Effectively this increases the direct conductivity of the medium, reduces  $E$  and consequently reduces the velocity of the stream. It can readily be shown that the effect of the polarization field is negligible when  $\Omega^2\tau^2 \gg 1$ , and is small when  $\Omega\tau > 2$ . But if  $\theta = 3 \times 10^4$  °K., then  $\Omega\tau = 1.7 \times 10^9 H/N$ , so that the condition  $\Omega\tau > 2$  becomes  $H/N > 1.2 \times 10^{-9}$ ; for  $N = 10^8$  or  $10^9$  per cc.,  $H > 0.12$  or  $1.2$  gauss respectively. Referring to the table, it is seen that transverse streams at elevations corresponding to these concentrations and satisfying the above conditions can approach fairly close to the neutral point. On the other hand the width of the stream is narrow where  $H$  is large, so that it is unlikely that observable features could occur close to the sunspot. We may note that when  $N$  considerably exceeds  $10^{10}$  per cc. the stream has large velocities only in regions so narrow as to be of negligible significance.

An upward stream of particles persists only as long as an adequate supply of matter is available to replace that already ejected. While we do not investigate the question of continuity, it seems to offer no difficulties; the pressure gradients set up seem quite sufficient to ensure replacement of matter in the low regions of the chromosphere.

## VII. THE VISIBILITY OF A STREAM OF PARTICLES

The visibility of a stream of particles depends, among other things, on the temperature and on the concentration of particles in the stream.

If we look at a great thickness of hydrogen in which the physical conditions are uniform, the central intensity of  $H\alpha$  is the same as that in a black body at the same "effective" temperature. This effective temperature is known for a range of physical conditions(5).

If the gas be at a kinetic temperature  $\theta$  of  $2.5 \times 10^4$  °K. and be irradiated by hemispherical radiation from a black body at temperature  $5 \times 10^3$  °K., the effective temperature varies with electron concentration, being  $5.37 \times 10^3$  °K. for  $N_e = 10^{13}$  per cc., but dropping to  $3.82 \times 10^3$  °K. for  $N_e = 10^{12}$ , and to  $3.35 \times 10^3$  °K. for  $N_e = 10^{10}$  °K. (Table 2). Thus gas in the upper chromosphere would have a lower effective temperature than gas in the lower chromosphere; normally this is not significant as the optical thickness of the upper chromosphere is small. The thickness of a stream of the type described here may, however, be of the order of  $10^9$  cm. or more in the line of sight. The absorption coefficient is very sensitive to electron concentration and temperature. If  $N_e$  is of the order of  $3 \times 10^{10}$  and  $\theta$  about  $2.5 \times 10^4$  °K. in the stream, the absorption coefficient

is about  $5 \times 10^{-10}$  per cm. (see Table 3), so that the total absorption coefficient is of the order of unity. The brightness of the gas at the centre of H $\alpha$  in this case is roughly that of a black body at the effective temperature, e.g.  $3.4 \times 10^3$  °K., and this is slightly lower than the H $\alpha$  background in the normal chromosphere, estimated theoretically for similar assumed conditions of irradiation to be  $3.58 \times 10^3$  °K.(6). Thus in this case the stream appears as a region of H $\alpha$  absorption when viewed against the sun's disk. Since the effective temperature drops as the electron concentration drops, it will generally be the case that such streams will appear as dark markings when viewed in H $\alpha$ ; the contrast will be greater if the absorption maximum be displaced by a Doppler shift. Conversely the stream would appear as a prominence if observed at the limb.

TABLE 2

EFFECTIVE TEMPERATURE OF HYDROGEN FOR THE EMISSION OF H $\alpha$   
GAS IRRADIATED BY HEMISPHERICAL BLACK-BODY RADIATION,  $T = 5 \times 10^3$  °K.

Kinetic Temperature of Hydrogen (°K.)	Effective Temperature (divided by $10^3$ )				
	Electron Concentration (per cc.)				
	$10^{13}$	$10^{12}$	$10^{11}$	$10^{10}$	$10^9$
$5.0 \times 10^4$	6.34	4.25	3.61	3.51	3.50
$2.5 \times 10^4$	5.37	3.82	3.42	3.35	3.35
$1.5 \times 10^4$	4.61	3.54	3.26	3.23	3.23

### VIII. TRANSVERSE STREAMS AND OBSERVED SOLAR PHENOMENA

The theory has been developed above for a simple bipolar pair of sunspots whose line of centres is perpendicular to the sun's general magnetic field, the latter being inclined to the sun's surface. It is clear, however, that the magnetic field due to other neighbouring sunspots may play a role similar to that of the sun's general field.

The trajectories of some transverse streams near a simple bipolar pair of sunspots are shown in Figure 1. In this case, the trajectories lie close to the dividing plane and some would appear as approximate arcs of ellipses. Such streams would be observable only if they contained sufficient neutral atoms; thus in general they are observable in the coronal regions only if they have risen from the chromosphere.

The stream slows down if it moves into a region of greater magnetic field strength, or if it is projected into a region where the electric field is weaker. Cowling has pointed out(7) that in the absence of an electric field a moving mass of gas would be retarded in crossing a magnetic field, its charges being dispersed along the lines of magnetic force. Thus a mass of gas denser than its surroundings



and falling under the action of gravity would have a comparatively small velocity across lines of magnetic force, but the component of the downward motion along the magnetic field would be unaffected, so that the direction of fall would coincide approximately with lines of magnetic force. The streams described above will usually return to the sun in this manner.

The velocities and particle concentrations predicted for transverse streams are of the order of those found in prominences, and they are expected, on the basis of this theory, to rise from localized regions in the upper chromosphere. The velocity is highest when the rate of change of flux through the sunspot is highest. It is just at such times, however, that chromospheric flares are expected, and thus there should be a close correlation between flares and transverse streams.

TABLE 3

THE ABSORPTION COEFFICIENT OF HYDROGEN AT THE CENTRE OF  $H\alpha$   
GAS IRRADIATED BY HEMISPHERICAL BLACK-BODY RADIATION,  $T=5 \times 10^3$  °K.

Kinetic Temperature of Hydrogen (°K.)	Absorption Coefficient (per cm.)				
	Electron Concentration (per cc.)				
	$10^{13}$	$10^{12}$	$10^{11}$	$10^{10}$	$10^9$
$5 \cdot 0 \times 10^4$	$1 \cdot 7 \times 10^{-7}$	$1 \cdot 3 \times 10^{-8}$	$4 \cdot 6 \times 10^{-10}$	$6 \cdot 2 \times 10^{-12}$	$6 \cdot 4 \times 10^{-14}$
$2 \cdot 5 \times 10^4$	$1 \cdot 3 \times 10^{-6}$	$1 \cdot 1 \times 10^{-7}$	$4 \cdot 0 \times 10^{-9}$	$5 \cdot 4 \times 10^{-11}$	$5 \cdot 6 \times 10^{-13}$
$1 \cdot 5 \times 10^4$	$9 \cdot 7 \times 10^{-6}$	$7 \cdot 5 \times 10^{-7}$	$2 \cdot 6 \times 10^{-8}$	$3 \cdot 4 \times 10^{-10}$	$3 \cdot 6 \times 10^{-12}$

We identify these streams with the class of eruptive prominences associated with flares, previously attributed to radiation pressure. Whilst radiation pressure undoubtedly acts on these prominences, its effects may be much less significant than hitherto believed.

### IX. ACKNOWLEDGMENT

The work described in this paper was carried out as part of the research programme of the Division of Physics, C.S.I.R.

### X. REFERENCES

- (1) GIOVANELLI, R. G.—*Mon. Not. Roy. Astr. Soc.* **107**: 338 (1947).
- (2) TOWNSEND, J. S.—*Proc. Roy. Soc. A* **86**: 571 (1912).
- (3) GIOVANELLI, R. G.—*Phil. Mag.* **40**: 206 (1949).
- (4) JEANS, J. H.—“Electricity and Magnetism.” 5th Ed., p. 268. (Cambridge Univ. Press, 1933.)
- (5) GIOVANELLI, R. G.—*Aust. J. Sci. Res. A* **1**: 305 (1948).
- (6) GIOVANELLI, R. G.—*Ibid.* **1**: 360 (1948).
- (7) COWLING, T. G.—*Proc. Roy. Soc. A* **183**: 453 (1945).

# SOLAR RADIATION AT 1200 MC/S., 600 MC/S., AND 200 MC/S.

By F. J. LEHANY\* and D. E. YABSLEY\*

(Plates 1-2)

[*Manuscript received September 9, 1948*]

## *Summary*

Daily observations of solar radiation at frequencies of 1200 Mc/s., 600 Mc/s., and 200 Mc/s. taken between August 18 and November 30, 1947, are described. The characteristics of the radiation at 200 Mc/s. were in general agreement with those observed by earlier workers. At 600 Mc/s. and 1200 Mc/s., the received intensity was normally steady on any one day but underwent long-period variations over a range of about two to one. The radiation received when the sun was almost free of sunspots corresponded to an effective black-body temperature of 0.5 million °K. at 600 Mc/s. and 0.1 million °K. at 1200 Mc/s. As sunspots appeared, the temperature rose and showed marked correlation with sunspot area. It is considered that radiation at these frequencies is entirely thermal in origin and that the long-period variations are at least partly due to the influence of the magnetic field of sunspots on the mechanism of thermal emission from a magneto-ionic medium.

On a few occasions, isolated disturbances were observed on 600 Mc/s. and 1200 Mc/s. some of which were associated with chromospheric flares and radio fade-outs.

The difficulties arising in the calibration of the apparatus and the steps taken to overcome them are discussed in detail.

## I. INTRODUCTION

The discovery(1, 2, 3, 4) that the effective black-body temperature of the sun, when observed at 200 Mc/s. and lower frequencies was usually one million °K. and often rose to several million degrees, was in sharp contrast with the results obtained by Southworth and others(5, 6, 7) at 10,000 Mc/s. and higher. The temperature at the latter frequencies was of the same order as the optical temperature and did not exhibit large variations in level.

In the presence of certain sunspots, especially during their central meridian passage, it was found that the radiation at low frequencies (200 Mc/s. and lower) was increased severalfold and, in addition, fluctuated violently in level. By accurate direction-finding, McCreedy, Pawsey, and Payne-Scott(2) showed that this enhanced radiation came from a region restricted to the immediate neighbourhood of the spot group concerned. At the higher frequencies, on the other hand, the meridian transit of sunspots did not appear to have any special significance.

This contrast in solar behaviour suggested that a study of the sun's radiation at intermediate frequencies would be of special interest. With this in view, the

\* Division of Radiophysics, C.S.I.R.

frequencies chosen for the present observations were 1200 Mc/s. and 600 Mc/s., together with 200 Mc/s. as an index of solar activity at the lower frequencies.

## II. EQUIPMENT

The aerial consisted of a 16 ft. by 18 ft. section of a paraboloid of revolution, and was used for the simultaneous reception of all three frequencies by using three separate feeds at the focus. To avoid undue signal loss at 600 Mc/s. and 1200 Mc/s., the lengths of the connecting cables were reduced to a minimum by mounting the first stages of the receivers on the aerial framework. Cable attenuation at 200 Mc/s. was not so serious and the receiver was installed in a nearby cubicle. A general view of the installation is shown in Plate 1 and of the aerial feed system in Plate 2.

The receivers were of normal radar design but with specially selected valves and crystals to give the best possible noise factors. At 200 Mc/s. adequate signal gain and selectivity were achieved with three stages of amplification using 6J4 grounded-grid triodes preceding the mixer and intermediate-frequency

TABLE 1

Frequency (Mc/s.)	Directivity (D) of Aerial Pattern	Beamwidth of Aerial Pattern to $\frac{1}{2}$ -Power Points (degrees)	Average Noise Factor of Receiver
200	104	21	7
600	1120	7	17
1200	4400	3.5	14

amplifier. At 600 and 1200 Mc/s. crystal converters were used and signal selectivity was achieved by the use of resonant cavities. The importance of adequate signal discrimination will be discussed in Section IV.

After amplification at the intermediate frequency of 30 Mc/s., the signals were detected and fed to three recording milliammeters. A summary of the more important equipment constants is given in Table 1.

## III. METHOD OF MAKING OBSERVATIONS

The paraboloid and its mount were originally part of a radar set and could not be readily adapted to follow the sun continuously. Both the declination and the axis of rotation, however, could be adjusted so that the paths of the sun and aerial across the sky intersected at two points symmetrical in time about true noon. Records were taken by turning the aerial just ahead of the sun and then letting the sun drift through the aerial pattern. After directing the aerial at the sky alone to check the receiver stability, it was again set just ahead of the sun, resulting in the type of record shown in Figure 1. The maximum values



of the envelope represent the points of coincidence of the sun and maximum aerial directivity. This value was used to calculate the solar temperature for that day. The rest of the record, usually about two hours in duration, could be used to observe the general behaviour of the radiation and in particular any sudden changes in level.

The noise factors of the receivers were measured before and after each record and the effective temperature of the sky was determined by replacing the aerial with a resistive load at ambient temperature. From these three observations the effective black-body temperature of the sun could be obtained as described in Appendix I.

#### IV. CALIBRATION AND SOURCES OF ERROR

The effective black-body temperature of a radiating source such as the sun may be obtained from measurements of the received power as shown in

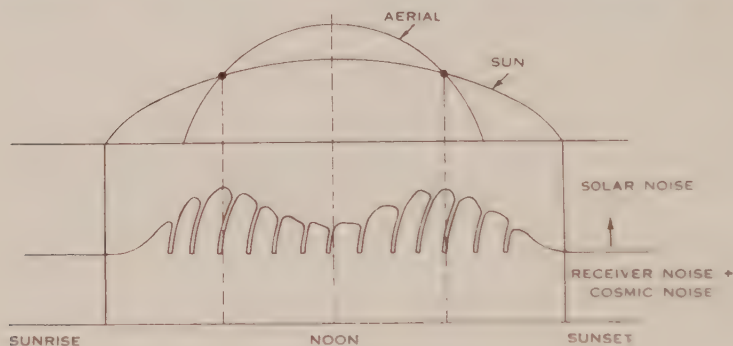


Fig. 1

Appendix I. It will be seen that this involves a knowledge of two instrumental constants, the receiver noise factor and the aerial directivity. Possible sources of error in determining these constants will now be discussed.

##### (a) Receiver Noise Factor

The receiver noise factor was determined by replacing the aerial by noise generators of known performance as described in Appendix II.

The basic assumption made in the calibration and use of each receiver is that the aerial circuit and noise generators are electrically interchangeable in all respects other than their noise production or effective temperature. In particular, their impedances should be as nearly identical as possible. Variations in impedance give rise to two sources of error. Firstly, the power transfer efficiency from the aerial to the receiver is dependent on impedance. Secondly, the receiver gain may alter as the input impedance is changed.

##### (b) Power Transfer Efficiency

At these frequencies it is, in general, difficult to adjust and maintain the impedances of the aerials and noise generators to produce a voltage standing

wave ratio of less than 1.1, mainly owing to the limitations of slotted lines and cable connectors. At 600 Mc/s. and 1200 Mc/s. the input circuits of the receivers in their final form were matched also to the cable impedance, and under this condition an impedance uncertainty of 10 per cent. causes insignificant variations in power transfer. At 200 Mc/s., however, to obtain optimum noise factor the receiver input impedance corresponded to a voltage standing wave ratio of 3 to 1. Under this condition an impedance change of 10 per cent. is equivalent to a power transfer variation of 6 per cent. This alone would cause the noise factor to vary over a range of approximately twice this amount as the length of connecting cable between aerial and receiver is varied.

### *(c) Gain Variations*

It was found that gain variations were likely to give rise to even more serious calibration errors. In the present equipment, the legitimate increases in power output due to the sun or the noise generators varied between 20 and 100 per cent. Any other change in output, such as a gain variation occurring at the same time as the input circuits are changed, cannot be separated from the legitimate power increase. A 10 per cent. change in gain could cause a 50 per cent. error in the worst case cited above. In the 200 Mc/s. receiver, considerable amplification occurs at the signal frequency and the presence of regeneration causes the gain to be sensitive to changes in input impedance. This effect was reduced to a minimum by using grounded-grid triodes and by careful shielding. The residual effect was then treated by an averaging process discussed below.

At 600 Mc/s. and 1200 Mc/s. no signal amplification was used. Initially this gave rise to serious calibration difficulties owing to the interaction of the signal input and local oscillator circuits through the crystal mixer. The difference in impedance between the aerial and noise generators at the frequency of the local oscillator resulted in different crystal operating conditions and rendered the calibration invalid. In addition, the receivers suffered from the usual sources of interference associated with superheterodyne reception without signal selectivity. Power was being received from both the sun and the noise generators at the image frequency and higher frequency bands using harmonics of the local oscillator for frequency conversion. Quite apart from calibration difficulties, the fact that the receivers were not effectively monochromatic would have detracted seriously from the value of any observations obtained with them.

The simplest and most effective cure for both these defects was to provide signal selectivity with a transmission type cavity resonator consisting of a quarter wavelength of concentric line. The unloaded  $Q$  of the cavities was of the order of 2000. Coupling was increased until the  $Q$  dropped to about 200 with a corresponding signal loss of one decibel. The coupling and tuning controls of the cavity provided two convenient degrees of freedom for adjusting the receiver input-impedance to the matched condition. In addition to eliminating unwanted bands of reception, the selectivity of the cavities also adequately protected the

local oscillator from the aerial and noise generator circuits, the operating conditions of the crystal mixer remaining constant as the latter two were interchanged.

#### *(d) Experimental Check of Calibration Errors*

A convenient and searching test for the types of error discussed above is to observe the receiver output while the length of cable between the receiver and the aerial or noise generator is varied. In general, cyclic output variations will occur as the cable length is altered. This test directly indicates the limits of the above errors associated with a particular aerial or noise generator. At 600 Mc/s. and 1200 Mc/s. a piece of telescopic line was permanently built into the equipment. At 200 Mc/s., selected lengths of line could be inserted. The variations in output were reduced to a minimum by impedance adjustment and then the residual variations were averaged to obtain a mean value of the output. This procedure was vital to the accurate comparison of receiver power outputs when, for instance, an aerial and a noise generator were interchanged. At 600 Mc/s. and 1200 Mc/s. the variations were negligible but at 200 Mc/s. they were significant and careful averaging was necessary.

A minor criticism of this variable-length line test is that an absence of variations indicates that either the generator or the receiver is matched to the line but not necessarily that both are matched. Thus in the special case of a matched receiver with no regeneration, any generator impedance will give a constant output and the test is no longer a check on the interchangeability of aerials and noise generators. This exception, however, is adequately covered by the approximate measurement of the impedances because it is just this case in which the power transfer, and so the noise factor, is least sensitive to impedance variations.

#### *(e) Sky Temperature*

It will be noted that in the method of making observations, the aerial was turned from the sky plus the sun to the sky alone. No change in aerial circuit was made and the difficulties and errors associated with such a change were avoided. The sky temperature itself, nevertheless, had to be determined to complete the measurement and this necessitated replacing the aerial with a resistive load at ambient temperature and carrying out the averaging process explained above. It was not, however, necessary to make this comparison very often as the sky temperature was sufficiently constant with time. The ratio of power outputs  $P_2/P_1$ , which is the factor directly concerned in equation (5) of Appendix I, did not at any time differ from unity by more than 2 per cent. for the part of the sky used in the present observations. It might be pointed out that this region was well clear of the Milky Way which is the major source of cosmic noise at 200 Mc/s. and lower frequencies.

Another assumption implicit in the above technique is that the temperature of the background sky surrounding the sun ( $T_b$ ) is equal to that of the reference part of the sky when the aerial is turned away from the sun. Although this condition is adequately satisfied from August to November, it is by no means



the case when the sun is in the direction of the Milky Way. In such circumstances the present technique is not applicable and  $T_E$  must be determined by a separate survey of each part of the sky when the sun is not present.

#### (f) *Aerial Directivity*

At present there is no wholly satisfactory method of measuring aerial directivity. The most desirable method is direct comparison with a standard aerial such as a half-wave dipole, but an accurate measurement is difficult to make as reflections from the earth and other bodies must be avoided. In the present case, no suitable opportunity of measuring the directivity by this method has presented itself and the method used was to estimate all the losses which contribute to a decrease in directivity below the theoretical value calculated from the dimensions of the aerial aperture.

In determining the directivity of the aerial it is convenient to consider the case when it is transmitting rather than receiving power. By the law of reciprocity, the directivities in the two cases are identical.

An aperture of area  $A$  fed in uniform phase and amplitude from a point source from which all the radiated energy falls on the aperture, has a directivity of  $(4\pi A)/\lambda^2$  compared with an isotropic radiator. In practice several factors cause a reduction in directivity below this value. In the present case these factors are :

- (i) Loss due to portion of the energy radiated from the feed missing the reflector altogether.
- (ii) Non-uniform distribution of amplitude and phase across the aerial aperture.
- (iii) Blockage of portion of the aperture by the feed.
- (iv) Leakage of portion of the energy through the reflector, which consists of a series of parallel rods instead of a continuous sheet.

The loss due to the first factor can be calculated directly from the polar diagrams of the feed and the geometry of the system. The loss due to non-uniform illumination of the aperture was calculated from curves given by Ryan(8). The loss due to blockage was assumed to equal the fraction of the aperture obscured by the feed. Finally, the loss due to leakage, which was very small in all cases, was calculated from curves given by Gooden(9). These factors have been collected in Table 2 together with estimated directivity values derived from them.

Ryan also gives the beamwidth of any aperture as a function of the illumination and this provides some check on the performance of the aerial as a whole. The theoretical total widths of the beam to the quarter power points are 3.5, 7.0, and 21 degrees for 1200, 600, and 200 Mc/s. respectively, for the plane in which the aerial width is 18 ft. These beamwidths were measured using the sun as source, and the values obtained were in good agreement with those given above.

The above method of estimating aerial directivity should give values representing the upper limit possible for the particular aerial arrangement. Any

error would mean that the values of intensity recorded in this paper set a lower limit to the true values.

### (g) Accuracy of Observations

It is estimated that the accuracy of the temperature and flux density results plotted in Figure 3 is of the order of 20 per cent. for absolute values and 10 per cent. for relative values.

## V. DISCUSSION OF OBSERVATIONS

### (a) 200 Mc/s. Records

The records on 200 Mc/s. showed the same general characteristics as have been described by McCready, Pawsey, and Payne-Scott(2) and other workers in this frequency range. On "quiet" days the received intensity was comparatively constant for the duration of the records and corresponded to a black-body temperature of one million °K. Occasional bursts\* of radiation were

TABLE 2

Frequency (Mc/s.)	Reduction in Directivity due to				Directivity		
	Energy Missing Reflector "w"	Non- uniform Illumina- tion "x"	Blockage by Feed "y"	Leakage through Reflector "z"	Products $wxyz = k$	$\frac{4\pi A}{\lambda^2}$	Estimated Directivity $= k \cdot \frac{4\pi A}{\lambda^2}$
200	0.75	0.98	0.96	0.98	0.69	150	104
600	0.91	0.95	0.96	1.00	0.83	1350	1120
1200	0.90	0.94	0.96	1.00	0.81	5400	4400

superimposed on the general level. On days of increased activity, which were usually associated with the central meridian passage of a sunspot group, the general level of radiation increased severalfold and in addition there was an almost continuous succession of bursts. McCready, Pawsey, and Payne-Scott(2) have shown that this increased radiation comes from the immediate vicinity of the sunspot group concerned. If this radiation were thermal in origin it would require this comparatively small region of the sun's atmosphere to have a temperature of several thousand million degrees. It is much more likely that a separate non-thermal source is responsible for this increase.

### (b) 600 Mc/s. and 1200 Mc/s. Records

The records on 600 Mc/s. and 1200 Mc/s. were similar to each other in general character and may be conveniently described together. The level of radiation

\* A sudden increase in radiation lasting for a few seconds has been regarded as a burst.

on any one day was fairly constant and bursts were rare. The radiation received when the hemisphere facing the earth was almost free of sunspots corresponded to a black-body temperature of 0.5 million  $^{\circ}\text{K}$ . at 600 Mc/s. and 0.1 million  $^{\circ}\text{K}$ . at 1200 Mc/s. These values are in good agreement with those estimated by Martyn(10) in his theoretical treatment of thermal radiation from the quiet sun. As sunspots appeared the temperatures rose gradually; the maximum steady levels recorded were 0.9 million  $^{\circ}\text{K}$ . at 600 Mc/s. and 0.25 million  $^{\circ}\text{K}$ . at 1200 Mc/s.

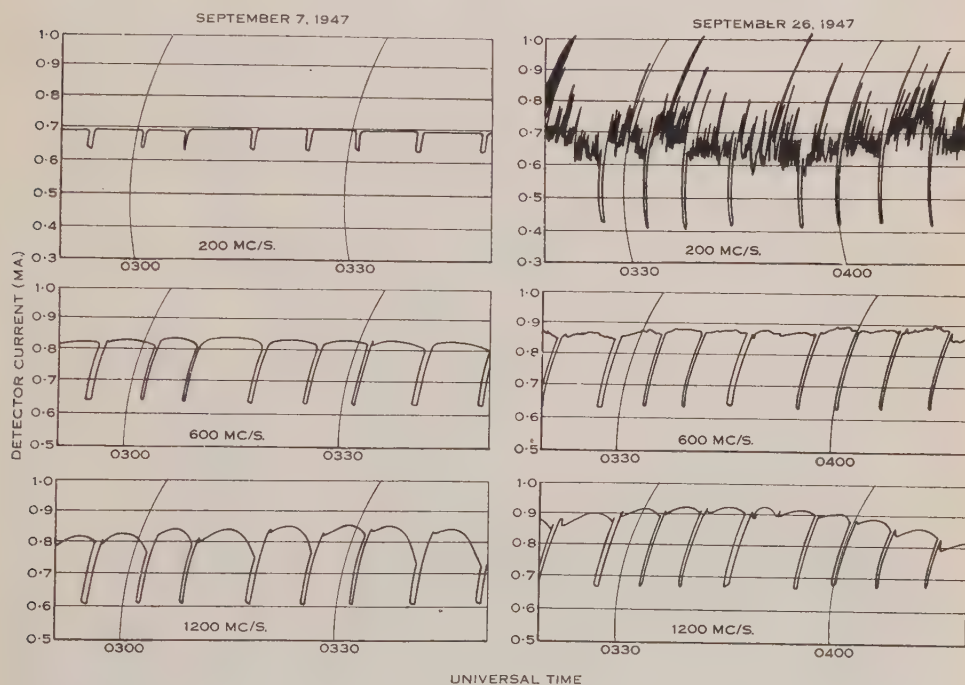


Fig. 2

The contrast in behaviour at 200 Mc/s. and these frequencies is illustrated by the typical records given in Figure 2. September 7 has been chosen as a representative day of low activity at 200 Mc/s. and September 26 as a day of high activity. The 600 Mc/s. and 1200 Mc/s. records do not reveal any major differences between the two days.

(c) *Evidence of Small Bursts at 600 Mc/s.*

It will be seen, however, that corresponding to the considerable number of bursts on 200 Mc/s. there was an unsteady trace on the 600 Mc/s. record suggesting that relatively small bursts were also occurring there. An unsteady trace of this nature occurred on September 25–29 inclusive. It will be noted from Figure 3 that this was also a period of intense solar activity on 200 Mc/s. These 600 Mc/s. fluctuations, if solar in origin, give some indication of the upper frequency limit of the variable non-thermal component so typical of the active sun on low



frequencies. It is estimated that the fluctuations represent a mean intensity of  $2.5 \times 10^{-22}$  watts. m.<sup>-2</sup>(c/s.)<sup>-1</sup>. The variable component on the corresponding 200 Mc/s. record represents a mean intensity of  $3 \times 10^{-21}$  watts. m.<sup>-2</sup>(c.s.)<sup>-1</sup>.

(d) *Variation of 600 Mc/s. and 1200 Mc/s. Radiation with Sunspot Area*

The daily values of the effective solar black-body temperature for the three frequencies are plotted in Figure 3 together with the projected sunspot area for

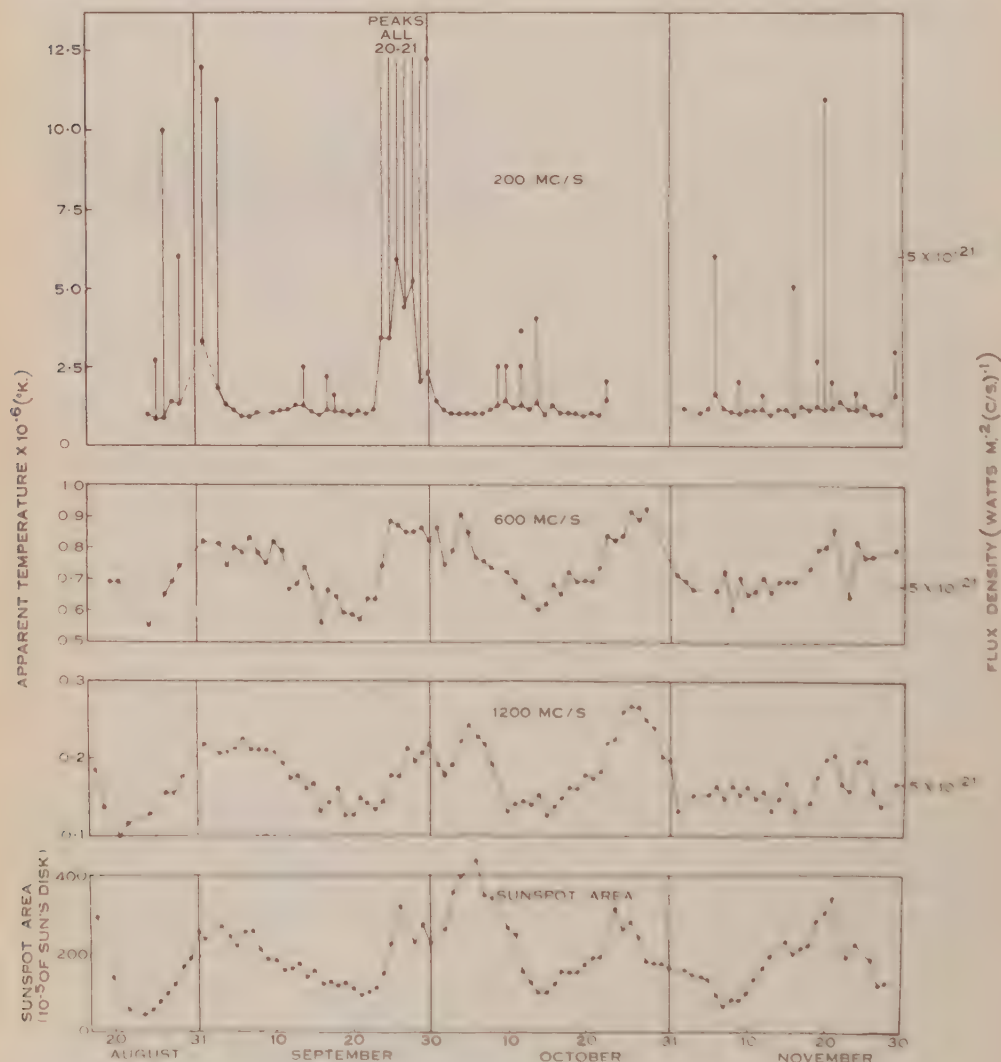


Fig. 3

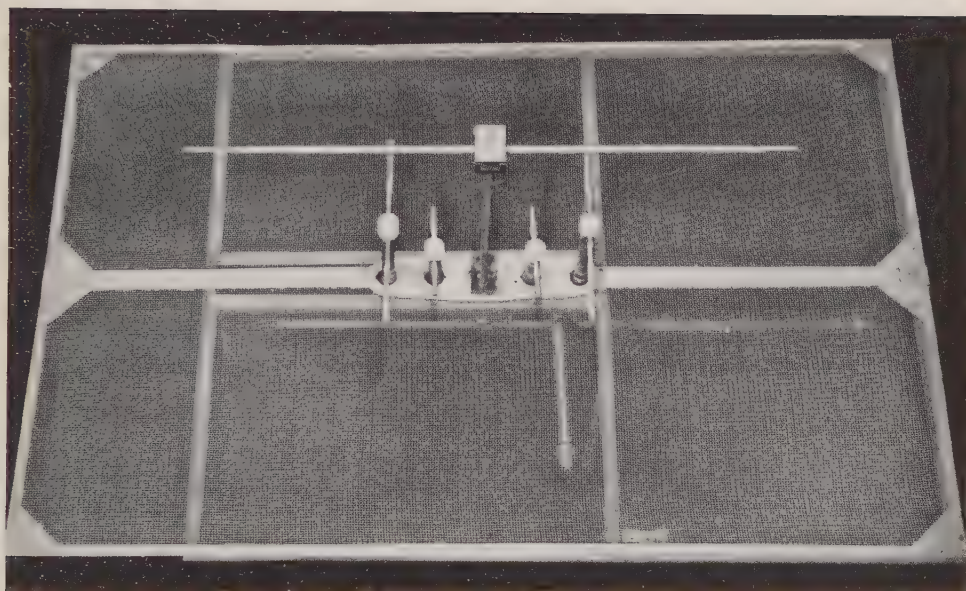
the period August 18–November 30, 1947. One of the more striking features of this figure is the correlation between sunspot area and solar radiation at 600 Mc/s. and 1200 Mc/s. This is particularly marked in the first month or two of observations but diminishes somewhat towards the end of the series.



LEHANY AND YABSLEY.—SOLAR RADIATION AT 1200 Mc/s., 600 Mc/s., AND 200 Mc/s.







LEHANY AND YABSLEY.—SOLAR RADIATION AT 1200 Mc/s., 600 Mc/s., AND 200 Mc/s.



It is thought that, apart from some isolated exceptions to be discussed later, all radiation at these frequencies is thermal in origin and these long-period variations in level are at least partly due to the influence of the magnetic field of sunspots on the effective levels in the solar atmosphere from which thermal radiation of a particular frequency band is produced. Some idea of this mechanism may be gained from the following considerations.

Magneto-ionic theory applied to the solar atmosphere shows that for a given radio frequency there exist two levels in the atmosphere at which the refractive index  $\mu$  becomes zero, one each for the "ordinary" and "extraordinary" rays. Further in towards the centre of the sun the electron density is greater than this critical value and the refractive index is imaginary. Radiation cannot be propagated when  $\mu=0$  and consequently can escape only if it originates in the less dense atmosphere outside the  $\mu=0$  layer. The amount of thermal radiation produced depends on both the temperature and emissivity of this atmosphere at the frequency under consideration and has been discussed in detail by Martyn(10). For simplicity it is assumed here that the temperature of the  $\mu=0$  level is an index of the radiation emitted.

From data on the sun's atmosphere compiled by Smerd(11) it appears that both the "ordinary" and "extraordinary"  $\mu=0$  levels for 600 Mc/s. and 1200 Mc/s. are normally in the chromosphere (temperature 30,000 °K.).\*

The much larger magnetic field of a sunspot raises the "extraordinary" level in this vicinity into the corona (temperature one million °K.) with a corresponding increase in the "extraordinary" component of radiation from that region. A quantitative check of this mechanism, however, would have to include a detailed consideration of each sunspot taking into account both its position on the sun and the strength of its magnetic field instead of the integrated sunspot area.

At 200 Mc/s. the  $\mu=0$  levels are normally in the corona and the effect of the sunspot magnetic field would push the "extraordinary" shell further into the corona. The resulting temperature variation would be small and is probably overshadowed by the other sources of non-thermal radiation shown to be associated with sunspots at the lower frequencies.

#### *(e) Central Meridian Passage of Sunspots*

Apart from the small bursts at 600 Mc/s., it appears that in contrast to 200 Mc/s. the central meridian passage of sunspots has no special significance at 600 Mc/s. and 1200 Mc/s. other than that inherent in the correlation with sunspot area. Observations by Covington(12) show that this is also true at 2800 Mc/s.

\* This would suggest that the quiet sun should have a temperature of 30,000 °K. This simple concept, however, deals only with radiation leaving the sun in a radial direction and applies in practice to the central portion of the sun's disk. Martyn(10) has pointed out that at these frequencies much of the radiation received from the edge of the sun's disk comes from the corona. It is presumably this limb-brightening effect at these frequencies which results in effective temperatures for the quiet sun considerably in excess of the temperature of the chromosphere.

*(f) Isolated Disturbances at 600 Mc/s. and 1200 Mc/s.*

In order to present the main characteristics of 600 Mc/s. and 1200 Mc/s. radiation in a simple form no mention has been made of some specific exceptions to the above general behaviour.

Isolated disturbances on 600 Mc/s. and 1200 Mc/s. were observed and were mainly of low intensity and fairly short duration. In some instances they were definitely associated with chromospheric flares and sudden daylight radio fade-

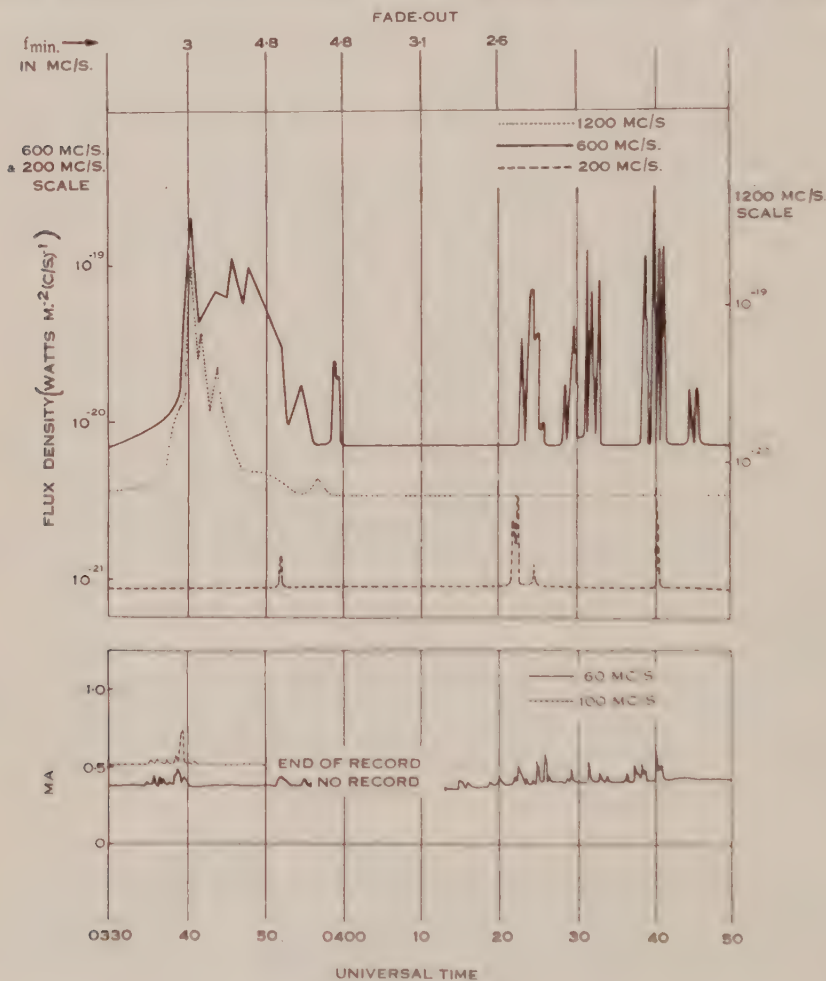


Fig. 4

outs. Sometimes no other solar observations were in progress and for this reason an exhaustive correlation is not possible. All the relevant details of the events, however, are summarized in Table 3.

The largest 600 Mc/s. and 1200 Mc/s. disturbance in the series occurred on October 4, when the intensity levels at both frequencies increased to thirty times their normal values and remained high for approximately ten minutes(13).



TABLE 3

A SUMMARY OF DISTURBANCES OBSERVED ON 600 Mc/s. AND 1200 Mc/s. TOGETHER WITH OTHER ASSOCIATED PHENOMENA (200 Mc/s. BURSTS, CHROMOSPHERIC FLARES, AND SUDDEN RADIO FADE-OUTS)

Date	Flares or Fade-outs	Frequency		
		1200 Mc/s.	600 Mc/s.	200 Mc/s.
1947				
Aug. 27	—	—	0512½ Burst 1.5	G.A. 3 to 4
29	—	—	0357-58 O.B. 1.3	G.A. ~3
Sept. 1	Fade-out	0355-0400 OB. 1.2	—	G.A. up to 3
3	Flare 2120	2121½ Burst 2	2121 Burst 4	2116-2121 Bursts 8
10	—	—	0407½-08 OB. 1.6	0405 and 0407 OB. s < 3
15	Fade-out	—	—	0252-53 OB. 10
19	—	0505-07 OB. 1.7	0504½-07 OB. 2	0504½-06 OB. > 3½
19	—	Not observing	0244-46 OB. 1.2	0243½-44½ OB. > 5
22	—	0427½-28 OB. 1.35	0427½-28 OB. 2.2	0427-27½ OB. 5
24	Flare	—	—	—
Oct. 2	2 Flares	—	—	Not observing
4	Fade-out	0333-0350 (approx.) Peak 0340 OB. 30	0330-0400 (approx.) Peak 0340 OB. 30	0352 Burst 1.5
4	„	—	0322½-0445 Series of Bursts Peak 40	0422 and 0440 2 Bursts 3
6	—	0317-19 Bursts 1.1	0316 and 0319 OB. 2 and 1.2	—
7	Fade-out	—	—	—
11	—	—	0324½ Burst 3.5	—
19	—	0244½-45½ OB. 1.05	0245-45½ OB. 6½	—
21	—	0201½ Burst 1.5	0201-03 OB. 1.5	0158-0202½ OB. > 5
21	Fade-out	—	0230½-35 OB. 1.5	—
21	—	—	0234½-35 OB. 1.5	—
22	—	—	0340-40½ OB. 3	—
Nov. 16	—	0246 Burst 1.3	—	0230-0250 G.A. 3
19	—	—	0130 and 0235 Bursts 1.5, 1.3	G.A. 3
20	Fade-out	—	0036½-38 OB. 1.15	0035-36 OB. 2
20	—	—	0310 and 0310½ Bursts 1.3	—
21	—	0109-0111 OB. 1.5	Not observing	G.A. 2
21	Flare and Fade-out	0144-0150 Peak 0147 OB. 1.5	0140-0154 Peak 0147 OB. 4	Fairly quiet
24	Fade-out	0111 Burst 1.15	—	—
„	„	0203½-05 OB. 2	—	—
„	„	0216-17 OB. < 1.05	0212½-18 OB. 1.5	—
25	Fade-out	—	—	—

Each disturbance is summarized by giving:

1. Time of occurrence and duration of the disturbance (universal time).

2. Classification of the disturbance under the following headings:

(a) Burst, being a sharp increase in intensity lasting for only a few seconds.

(b) Outburst (OB.), an increase in intensity lasting for a period of minutes.

(c) General activity (G.A.). This has been used to indicate the occasions on which disturbances on 600 and 1200 Mc/s. were accompanied by a continuous succession of bursts on 200 Mc/s., rendering impossible the identification of a corresponding event, if any, on this frequency. Otherwise, general activity on 200 Mc/s. has not been recorded in this table but has been indicated in Figure 3.

3. Peak intensity reached, relative to the prevailing steady level.

*Example.*—November 16. 1200 Mc/s.—at 0246 hours a burst occurred and reached a peak intensity of 1.3 times the existing level. 200 Mc/s.—general activity occurred from 0230 to 0250 hours with an intensity peak of 3 times the base level.

Such behaviour at lower frequencies is often associated with a chromospheric flare and may be conveniently termed an outburst. On this occasion a radio fade-out was reported by the Mt. Stromlo ionospheric recording station and is an almost certain indication of a flare. Details of the outburst and radio fade-out together with simultaneous records at 200 Mc/s., 100 Mc/s., and 60 Mc/s. are given in Figure 4. An interesting feature is the absence of an outburst at 200 Mc/s. At 100 Mc/s. and 60 Mc/s. a group of small bursts coincided with the commencement of the outburst at the higher frequencies.

## VI. CONCLUSIONS

The conclusions that may be drawn from this series of observations are:

- (i) The rapidly varying component of solar radiation which was observed at 200 Mc/s. and was associated with the central meridian passage of some sunspots had no observable counterpart at 1200 Mc/s. There was some evidence that such a component was occasionally present at 600 Mc/s. The variation had a mean intensity of approximately one-twelfth of the intensity of the corresponding variations at 200 Mc/s.
- (ii) The intensity of radiation at 600 Mc/s. and 1200 Mc/s. varied from day to day over an extreme range of about 2 to 1 and had marked correlation with sunspot area. The intensities observed are compatible with the thermal radiation expected from the solar atmosphere at these frequencies. It is considered that the variations in intensity with sunspot area are at least partly due to the magnetic fields of the sunspots raising parts of the effective radiating shells into the corona which is at a temperature of at least a million degrees.
- (iii) There were occasional disturbances on 600 Mc/s. and 1200 Mc/s., some of which were associated with chromospheric flares and radio fade outs.

## VII. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Radiophysics, C.S.I.R.

The authors wish to thank Dr. J. L. Pawsey for helpful discussions, and the staff of the Commonwealth Observatory, Mt. Stromlo, Canberra, for the data on solar flares and sunspot areas.

## VIII. REFERENCES

- (1) APPLETON, E. V., and HEY, J. S.—Solar radio noise—I. *Phil. Mag.* **37**: 73-84 (1946).
- (2) MCCREADY, L. L., PAWSEY, J. L., and PAYNE SCOTT, R. — Solar radiation at radio frequencies and its relation to sunspots. *Proc. Roy. Soc. A* **190**: 357-75 (1947).
- (3) PAWSEY, J. L. — Observations of million-degree thermal radiation from the sun at a wavelength of 1.5 metres. *Nature* **158**: 633-4 (1946).
- (4) RYLE, M., and VONBERG, D. D. — Solar radiation on 17.5 Mc/s. *Ibid.* **158**: 339-40 (1946).
- (5) SOUTHWORTH, G. — Microwave radiation from the sun. *J. Franklin Inst.* **239**: 285-97 (1945); correction—*Ibid.* **241** (1946).

- (6) DICKE, R. H., and BERINGER, R.—Microwave radiation from the sun and moon. *Astrophys. J.* **103** : 375-6 (1946).
- (7) SANDER, K. F.—Radio noise from the sun at 3.2 cm. *Nature* **159** : 506 (1947).
- (8) RYAN, J. P.—Design data relating directional diagram to aperture illumination of pencil and fan beam aeriels. Coun. Sci. Industr. Res. Aust. Radiophysics Lab. Rep. RP 241 (1945).
- (9) GOODEN, J. S.—Reflecting properties of metal gratings. *Ibid.* Rep. RP 215 (1944).
- (10) MARTYN, D. F.—Solar radiation in the radio spectrum—I. Radiation from the quiet sun. *Proc. Roy. Soc. A* **193** : 44-59 (1948).
- (11) SMERD, S. F.—The parameters of the sun's atmosphere in a form suitable for the application of the magneto-ionic theory. Coun. Sci. Industr. Res. Aust. Radiophysics Lab. Rep. RPL 14 (1948).
- (12) COVINGTON, A. E.—Solar noise observations on 10.7 cm. *Proc. Inst. Radio Engrs. N.Y.* **36** : 454-577 (1948).
- (13) LEHANY, F. J., and YABSLEY, D. E.—A solar noise outburst at 600 Mc/s. and 1200 Mc/s. *Nature* **161** : 645 (1948).

## EXPLANATION OF PLATES 1-2

### PLATE 1

General view of aerial and cubicle containing the 200 Mc/s. receiver together with local oscillators for the 600 Mc/s. and 1200 Mc/s. receivers.

### PLATE 2

Aerial feed system for 200 Mc/s., 600 Mc/s., and 1200 Mc/s.

## APPENDIX I

### *Determination of Effective Black-body Temperature of the Sun*

From the definition of the noise factor ( $N$ ) of a receiver it follows that, when the receiver aerial is replaced by an equal resistive impedance at ambient temperature, the power output is given by

$$P_1 = gkB_N(T_R + (N-1)T_R) \dots\dots\dots (1)$$

where  $g$  is the power gain of the receiver referred to the maximum thermal fluctuation power available from the resistance ( $kT_R B_N$ ),

$T_R$  is the temperature of the resistance in °K.,

$B_N$  is the bandwidth of the receiver,

$k$  is Boltzmann's constant.

When the aerial is surrounded by a constant temperature enclosure, the temperature of the radiation resistance of the aerial is that of the enclosure. Thus, when the aerial is pointed towards the sky the power output of the receiver is given by

$$P_2 = gkB_N(T_E + (N-1)T_R) \dots\dots\dots (2)$$

where  $T_E$  is the effective black-body temperature of the sky at the particular frequency concerned.

Likewise, when the aerial is directed towards the sun with a sky back ground temperature of  $T_E$ , the power output is

$$P_3 = gkB_N \left( T_E + (T_S - T_E) D \frac{\Omega_s}{4\pi} + (N-1)T_R \right) \dots\dots\dots (3)$$

where  $T_S$  is the effective black-body temperature of the sun,

$D$  is the directivity of the aerial pattern in the direction of the sun.

$\Omega_s$  is the solid angle subtended by the sun.

The above three equations give rise to the relation

$$T_S - T_E = \left( \frac{P_3}{P_2} - 1 \right) \frac{P_2}{P_1} \cdot N \cdot T_R \cdot \frac{4\pi}{D\Omega_s} \dots\dots\dots (4)$$

In the present series of observations  $T_E$  was always less than 300°K. and the lowest value of  $T_S$  was 100,000°K. In practice (4) was approximated to

$$T_S = \left( \frac{P_3}{P_2} - 1 \right) \frac{P_2}{P_1} \cdot N \cdot T_R \cdot \frac{4\pi}{D\Omega_s} \dots\dots\dots (5)$$

To obtain the effective black-body temperature of the sun ( $T_S$ ) from (5), the two instrument constants which must be known are the receiver noise factor ( $N$ ) and the aerial directivity ( $D$ ).

## APPENDIX II

### *Determination of Receiver Noise Factor (N)*

From (1) and (2) we obtain the relation

$$\frac{P_1}{P_2} = \frac{NT_R}{T_E + (N-1)T_R} \dots\dots\dots (6)$$

If instead of altering the temperature of the aerial to  $T_E$  by pointing it at the sky, we replace the aerial by a resistance of known high temperature (or any other noise generator of known performance) then the noise factor  $N$  may be obtained by using (6).

In the present work, the noise generators for 600 Mc/s. and 1200 Mc/s. were the hot resistance type. Each consisted of a small straight-filament tungsten lamp mounted as the inner conductor of a coaxial transmission line together with two matching stubs. At 200 Mc/s. a temperature-limited diode noise generator was used.



# MICROWAVE THERMAL RADIATION FROM THE MOON

By J. H. PIDDINGTON\* and H. C. MINNETT\*

[*Manuscript received October 6, 1948*]

## *Summary*

Measurements have been made of the thermal radiation from the moon in a 15 Mc/s. band, centred at 24,000 Mc/s. The radiation from the whole lunar disk has been measured during several phase cycles. Temperatures corresponding to this radiation have been deduced and curves of temperature against phase angle drawn. One of these corresponds to average temperature over the disk and the other to the temperature of a point on the lunar equator. They are found to be approximately sinusoidal, with amplitude  $\pm 40.3$  °K. and  $\pm 52.0$  °K. respectively and with a phase lag behind the lunar phase angle of about  $45^\circ$  in each case.

These results conflict with previous measurements of temperature at long infra-red wavelengths in both amplitude and phase angle of the temperature curve. An explanation is given in terms of radiation from subsurface layers of the moon's crust, which are partially transparent to the electromagnetic waves with which we are dealing. The theory is developed quantitatively and it is found that the results are consistent with the existence of a thin layer of dust covering a solid lunar surface.

Estimates are made of the temperatures of the disk of the new moon (156 °K.) and of the deep interior (241 °K.).

## I. INTRODUCTION

For many years the temperature of the moon's surface has been studied by means of thermocouples responding to the heat energy penetrating the earth's atmosphere through a "window" in the 8-14 micron region of the spectrum. The war-time development of sensitive radio receivers for micro-wavelengths provides a new method for the measurement of thermal radiation. Dicke(1) has discussed the design of a receiver for this purpose and some measurements of the temperature of the sun and of the nearly full moon have been reported by Dicke and Beringer(2).

This paper describes measurements made with similar equipment sensitive to a narrow band of the spectrum at a wavelength of 1.25 centimetres (24,000 Mc/s.). During three lunar cycles between April and July 1948, thermal radiation from the moon's disk was measured and the results interpreted in terms of lunar temperature.

## II. PREVIOUS WORK

Pettit and Nicholson(3) and Pettit(4, 5) have made numerous observations of long infra-red radiation from the moon using thermocouples. They found that the intensity of radiation varies over the lunar cycle through a large range

\* Division of Radiophysics, C.S.I.R.

and that the curve of intensity is symmetrical and in phase with the lunar phase angle. The curve produced by Pettit(4) showing variation of the total planetary radiation is given in Figure 1. The radiation received varies from a value of 600 (arbitrary units) to zero.

In an earlier paper, Pettit and Nicholson(3), using a thermocouple small compared with the lunar image formed by the telescope, measured the radiation from the centre of the new moon. They found a temperature of  $120^{\circ}\text{K}$ . The quantity of radiation detected was small, being near the limit of sensitivity of their thermocouple, so that the accuracy of the measurement was low.

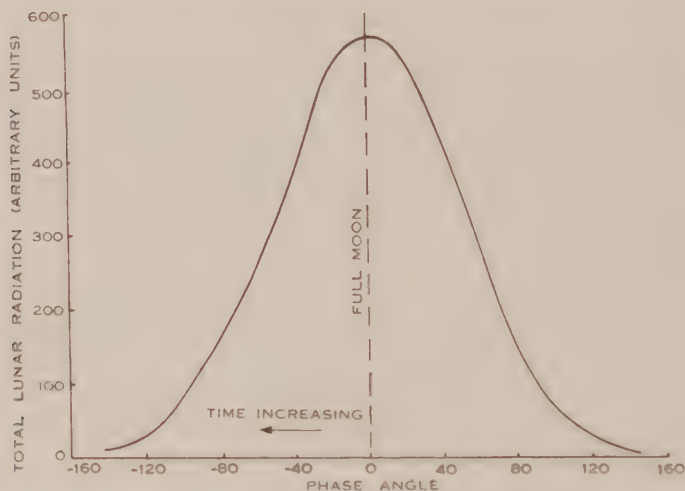


Fig. 1.—Lunar radiation in the 8-14 micron region (after Pettit).

Using the same apparatus, Pettit and Nicholson also measured the quantity of radiation  $E$  received from various parts of the full moon. They found the law of variation of  $E$  to be given by  $E = E_s \cos^{2.3} \varphi$ , where  $\varphi$  is the angle between the normal to the area considered and the direction of the sun\* and  $E_s$  is the radiation from the subsolar point. Once again radiation varies over a large range and is in phase with the lunar phase angle.

In comparing our results with those of the above workers we have taken a value of  $E_s$  corresponding to a temperature of  $374^{\circ}\text{K}$ . of the subsolar point. The theoretical value of temperature of the subsolar point ( $374^{\circ}\text{K}$ .) given by Pettit and Nicholson is chosen rather than their observed value ( $391^{\circ}\text{K}$ .), as the latter is probably in error owing to inaccuracies in the assumed transmission properties of the earth's atmosphere. Also a law of  $\cos \varphi$  instead of  $\cos^{2.3} \varphi$  is used since the experimental result is probably due to a directivity effect in the radiating properties of the lunar surface.

Of importance in analysing our results is the fact that the curve of surface temperature is in phase with that of incident radiation. This is clearly shown

\* They give the expression  $E_s \cos^{3/2} \varphi$  which does not fit their results and appears to be an error.

by Pettit and Nicholson's and Pettit's curves described above and is due to the fact that the proportion of incident radiation conducted into the lunar interior is small. The latter fact is deduced by Pettit and Nicholson from analysis of lunar eclipse results. It may also be deduced from the lack of phase lag of the surface temperature curve. The flux of heat into the interior leads the surface temperature in phase by  $\pi/4$  (see, for example, Carslaw and Jaeger, 6, p. 57), and in order to combine with the reradiated energy to give a resultant with zero phase lag, the magnitude of this flux must be small.

Since the results of Pettit and Nicholson and of Pettit are in substantial agreement with earlier work and sufficiently comprehensive for a thorough comparison with our own results, no reference will be made here to numerous earlier observations of lunar radiation.

The first microwave measurements on the moon were those obtained by Dicke and Beringer(2) at a wavelength of 1.25 centimetres. The temperature of the nearly full moon (phase +18°) was found to be 292 °K. Their result should be corrected by a factor of 0.93 to allow for the actual semi-diameter of the moon on the occasion of measurement, namely, 16' 37" rather than the assumed value of 16'. This reduces the average disk temperature to 270 °K.

### III. EXPERIMENTAL METHOD

A brief description of the apparatus and method is given in the present paper. The complete equipment, the calibrating procedure, and the observational techniques found necessary to make effective use of the apparatus in radio astronomy will be described more fully elsewhere.

#### (a) Apparatus

The aerial consists of a parabolic reflector 44 inches in diameter and equatorially mounted. It is fed by means of a carefully matched horn. Calibrated scales are fitted to the hour angle and declination axes, which are adjustable by hand-wheels operating through reduction gears. Accurate alignment of the radio beam, which has a width to half power points of  $\pm\frac{3}{8}^\circ$ , is made by means of a visual telescope attached to the reflector. The power gain of the aerial relative to an isotropic radiator is  $3.0 \times 10^4$  (44.8 decibels).

A block diagram of the receiving equipment is shown in Figure 2. Following the method of Dicke(1), the energy received by the waveguide horn at the reflector focus is compared with the thermal radiation from a disk of absorbing material at a fixed temperature. The comparison is made 25 times per second by rotating the disk at this speed and allowing it to enter the waveguide once in each revolution. The energy thus modulated is converted to an intermediate frequency of 30 Mc/s. by beating with the output of a klystron oscillator in a balanced crystal mixer. These units, together with a 30 Mc/s. preamplifier, are contained in a box mounted near the mirror focus. The preamplifier output is connected by coaxial cable to the remaining 30 Mc/s. amplifying stages and then demodulated by a diode rectifier. The total bandwidth of the amplifier is 15 Mc/s. including the image response. The 25 c/s. modulation frequency is amplified and converted to direct current for operation of a recording milliammeter by



mixing with a 25 c/s. signal from a contactor rotating in synchronism with the modulating disk. A response time of either 1 or 6 seconds is determined by means of low-pass filters following the synchronous rectifier.

The overall sensitivity of the apparatus is such that radiation equivalent to an aerial temperature change of  $1.2^{\circ}\text{K}$ . produces an output equal to the root mean square of the random fluctuations in receiver output. This corresponds to lunar temperature changes of about  $8^{\circ}\text{K}$ . Since radiation in the microwave spectrum obeys the Rayleigh-Jeans law and is proportional in intensity to the temperature of the source, very low lunar temperatures could be measured.

It may be seen that the radiometer is much more sensitive than the thermocouples of earlier workers when low temperatures are being measured. The lowest temperature of a body from which Pettit and Nicholson could detect radiation was of the order of  $100^{\circ}\text{K}$ .

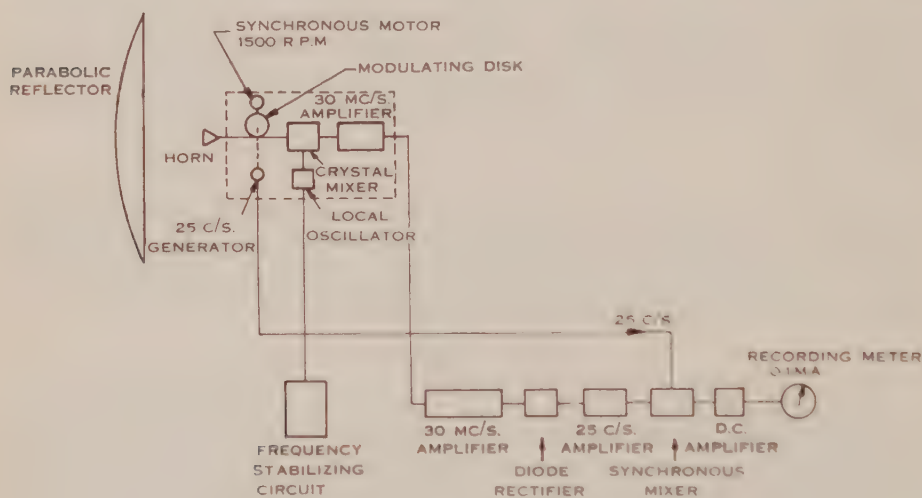


Fig. 2.—Block diagram of receiving equipment.

### (b) Observational Technique

Observations were made daily when the moon was at or near transit. The method was to set the aerial beam slightly ahead of the moon in hour angle and to record the receiver output as the moon passed through the beam. This procedure was repeated for a number of declination angles, as set by the visual telescope, until a maximum reading was established. Cosmic radio noise, radiation from the partially transparent earth's atmosphere and from the moon were all incident on the aerial during the observations. However, at these frequencies cosmic noise is negligible, and if the aerial beam is not moved during a measurement, the received atmospheric radiation is constant. The variation in receiver output is due, therefore, to the reception of lunar radiation as the moon passes through the beam.

As well as radiating themselves, the oxygen and water vapour of the earth's atmosphere will absorb part of the lunar radiation. The degree of absorption



was found by measuring the radiation from the atmosphere and a suitable correction was applied. The method used is a modification of the technique of Dicke *et al.*(7).

From the corrected value of the receiver output an equivalent aerial temperature was deduced by measuring the sensitivity of the receiver. This was done by replacing the aerial with a resistive termination which could be heated through a measured range of temperature. The corresponding change of receiver output provided the desired calibration. The temperature at which the termination would emit an amount of radiation equal to that received by the aerial is called the equivalent aerial temperature.

A further correction was necessary to allow for variations of the solid angle subtended by the moon. These values were taken from the *Nautical Almanac* and adjusted to allow for the fact that all observations were made at or near transit (i.e. at an average angle of  $35^\circ$  from the zenith). The figure obtained after these corrections were applied is the aerial temperature due to a moon of semi-diameter  $16'$ , in the absence of atmospheric absorption.

### (c) Reduction of Results

The interpretation of the measured aerial temperature in terms of lunar temperature depends on the directional properties of the aerial and the amount of microwave radiation from various parts of the moon's disk. Since our aerial beam has a width of only  $\pm 3^\circ$  it is evident that the aerial gain will differ for reception from different parts of the lunar disk. Radiation from its centre will be "amplified" more than that from the limb since the aerial gain is a maximum for a direction passing through the centre of the disk. Thus lunar temperatures calculated on the basis of maximum aerial gain are less than the true average disk temperatures. This factor is discussed in Appendix I together with the allowance which must be made for the fact that the lunar emissivity is less than unity. An average microwave emissivity of 0.9 is assumed for the whole lunar disk.

## IV. COMPARISON OF MICROWAVE AND INFRA-RED RESULTS

A series of lunar aerial temperature measurements was made during three phase cycles between April and July 1948. Some of these results were discarded because of unfavourable conditions due to cloud or strong winds. The remaining corrected aerial temperatures are plotted against phase angle in Figure 3. A sine wave which provides the best fit is also shown and may be represented by the equation

$$T = 39.0 + 6.6 \cos (\Omega t - 45^\circ) ^\circ\text{K.}$$

where  $\Omega$  is the angular velocity of the moon and  $t$  the time elapsed since full moon.\*

As described above, the variation in gain of the aerial beam over the moon's surface must be taken into account when determining the true space-average temperature over the lunar disk and the temperature of the spot on the lunar

\* Lunar phase angles decrease with increasing time.

equator nearest to the earth. Temperature curves of these two quantities corrected for emissivity are shown as curves *A* (continuous line) and *B* (broken line) in Figure 4. They may be represented by the equations

$$T \text{ (average)} = 239 + 40.3 \cos (\Omega t - 45^\circ),$$

$$T \text{ (equatorial)} = 249 + 52.0 \cos (\Omega t - 45^\circ).$$

The absolute values of  $T$  (average) and  $T$  (equatorial) are thought to be accurate within  $\pm 5$  per cent. Values of aerial temperature are consistent within themselves to a higher order of accuracy.

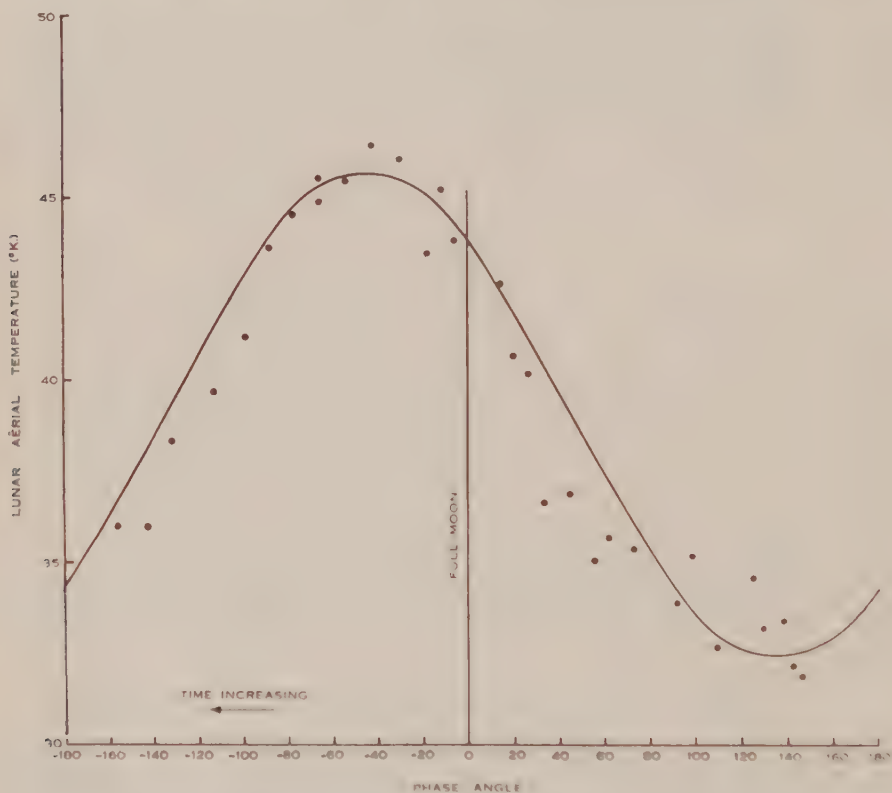


Fig. 3.—Observed lunar aerial temperature at 1.25 cm.

Also shown in Figure 4, as curve *C*, are Pettit's results for total radiation from the lunar disk. This is plotted in units of radiation received rather than temperatures and the two scales are related only at points corresponding to new and full moon. The curve shows, however, that there is no phase lag in surface temperature behind lunar phase. It is evident that if Pettit's curve of total radiation were redrawn in equivalent black-body temperature units, both in amplitude and phase it would differ markedly from our curve of average temperature over the disk. On the other hand, it will be noted that the single measurement made by Dicke and Beringer is in fair agreement with our results, when corrected for emissivity as above.

The apparent disagreement with the results of Pettit and of Pettit and Nicholson appears to be due to the fact that the material of the moon is partially transparent to 1.25-centimetre waves. Thus the radiation we detect originates in various layers beneath the surface, each layer contributing an amount depending on its opacity to this wavelength, its temperature, and the attenuation of radiation from it in reaching the surface. On the other hand, the lunar surface is substantially opaque to 8-14 micron radiation so that such measurements give surface temperatures. The surface temperature variation propagates

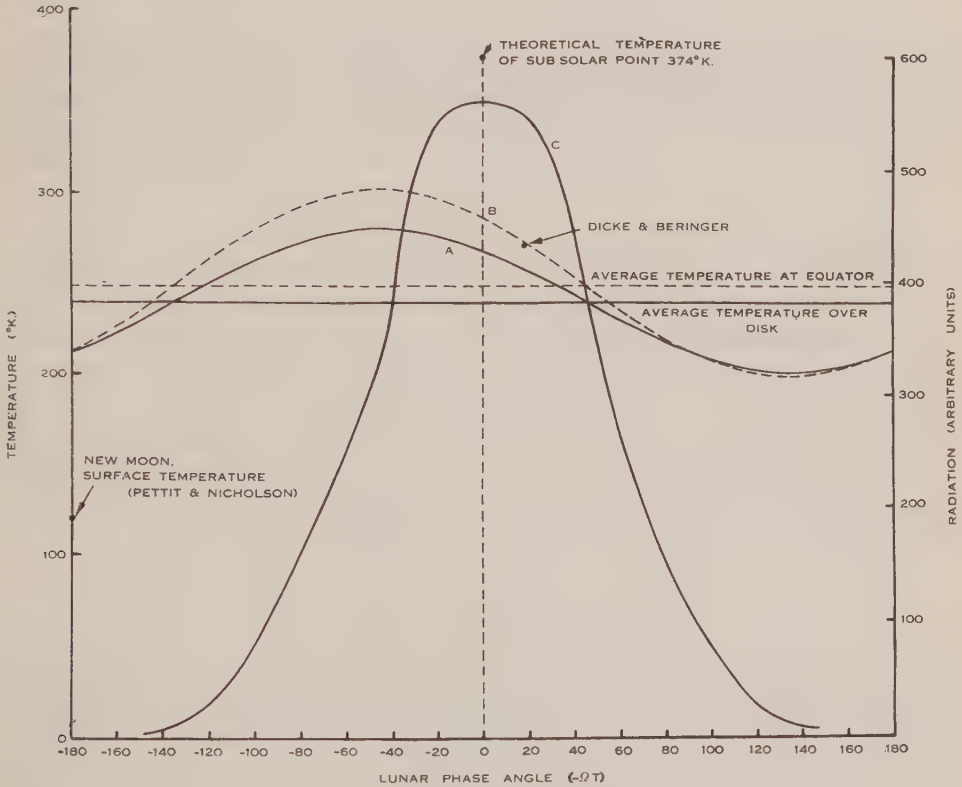


Fig. 4.—Lunar temperature curves.

- (A) Average microwave temperature over the moon's disk.
- (B) Microwave temperature of the centre of the disk.
- (C) Pettit's curve of radiation from the moon's disk.

into the moon as an attenuating temperature wave with a progressively lagging phase. Hence the integration of the radiation from all layers between the surface and the lowest layer from which radiation can reach the surface with an appreciable intensity leads to a "microwave" temperature which has a smaller periodic variation and which lags in phase behind the corresponding surface variation. These ideas are developed quantitatively in Section V.

We have also examined the possibility of discrepancy between the results due to the effect of reflection from the moon of 1.25-centimetre solar radiation

and the effect of an ionized lunar atmosphere. However, both are negligible, the first producing a change in apparent temperature of less than  $0.1^\circ\text{K}$ . The whole of the incident solar radiation would have to be absorbed in a layer of lunar atmosphere of thickness of the order of 1 centimetre if the resulting ionization were appreciably to affect propagation of microwave radiation. As the scale-height of a lunar atmosphere, if one exists, must be of the order of 50 kilometres, such a possibility need not be considered.

## V. INTERPRETATION OF RESULTS

Our results, shown in Figure 4, together with the estimate by Pettit and Nicholson of the subsolar surface temperature, allow us to make estimates of (a) the temperature far below the moon's surface, (b) the surface temperature of the dark side of the moon, and (c) the physical nature of the lunar surface.

### (a) *The Temperature of the Lunar Interior*

In the absence of any reservoir or source of heat in the lunar interior, the temperature some distance below the surface at the equator will be constant and equal to the steady component of equatorial temperature, namely,  $249^\circ\text{K}$ .

We have indicated in Appendix I that mean values of temperature over a lunar cycle both on and below the lunar surface are related to latitude by the equation  $T_\varphi = T_0 \cos^{\frac{1}{2}} \varphi$  where  $T_\varphi$  and  $T_0$  are the temperatures at latitudes  $\varphi$  and  $0$  respectively.  $T_0$  has a value  $249^\circ\text{K}$ . By using the equation for  $T_\varphi$  to integrate over the whole surface of the moon, we find the space-average of the steady component of temperature. This temperature,  $0.94 \times 249^\circ\text{K} = 234^\circ\text{K}$ , is our estimate of the temperature far below the surface of the moon.\*

### (b) *Minimum Surface Temperature*

The surface temperature of a point on the lunar equator during the period of illumination may be expressed as  $T_s \cos^{\frac{1}{2}} \lambda$ , where  $\lambda$  is the difference in longitude between the point considered and the subsolar point and the value of  $T_s$  is taken as  $374^\circ\text{K}$ . for reasons given above. For simplicity we assume that the temperature when not illuminated is constant and equal to  $T_m$ . The resultant curve of temperature against time is periodic and has a Fourier expansion  $T_0 + \sum T_n \cos n\lambda$ . The coefficients of the constant term and fundamental oscillation are given by

$$\begin{aligned} T_0 &= 0.44(T_s - T_m) + T_m, \\ T_1 &= 0.60(T_s - T_m). \end{aligned}$$

The constant term should be equal to the time-average of the microwave temperature of a point on the equator, namely,  $249^\circ\text{K}$ . If  $T_s$  is taken again as  $374^\circ\text{K}$ . we then have  $T_m = 151^\circ\text{K}$ . Using the  $\cos^{\frac{1}{2}} \varphi$  law to average over the lunar disk it may be shown that the space-average surface temperature of the whole disk of the new moon is  $0.96 \times 151^\circ\text{K} = 145^\circ\text{K}$ .

\* This temperature (an average over a sphere) differs from the steady component of temperature averaged over the lunar disk which is given above as  $239^\circ\text{K}$ .



The approximate nature of these results, owing to the assumption that surface cooling during lunar night is negligible, could be corrected if the amount of radiation cooling during this period were known.

Of importance in later analysis is the magnitude of the fundamental component of oscillation of surface temperature. This is given by  $T_1 = 0.6 (374 - 151) = 134$  °K. which may be compared with the value of 52.0 °K. for our microwave equatorial temperature fluctuation, the ratio of the two fluctuations being 2.58.

### (c) Lunar Temperature Waves

As explained above, the temperature of the moon's crust varies with depth. The microwave radiation originating in the subsurface layers and emerging through an element of the lunar surface will first be expressed in terms of the equivalent temperature of an isothermal black body.

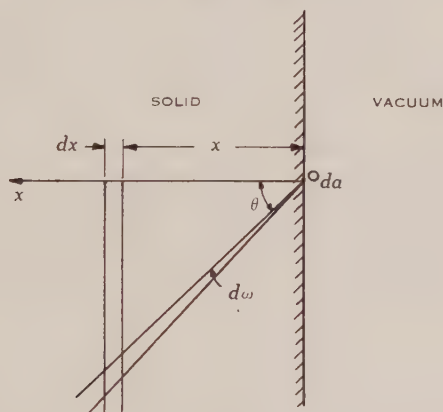


Fig. 5.—Radiation from a hot body.

Consider a radio wave of frequency  $f$  moving through the lunar material. If  $k$  is the volume absorption coefficient, the wave will attenuate according to the relation  $I = I_0 e^{-kx}$ , where  $I_0$  is the initial intensity\* and  $I$  the intensity after travelling a distance  $x$ . In dielectrics of low conductivity the value of  $k$  in c.g.s. units is  $4\pi\sigma/cK^{\frac{1}{2}}$ , where  $\sigma$  is the electrical conductivity,  $K$  the dielectric constant, and  $c$  the velocity of light. Also the refractive index  $\mu$  of the material is  $K^{\frac{1}{2}}$ . The necessary condition of low conductivity may be written  $2\sigma f/K \ll 1$ . Even for highly conducting dry soil having  $\sigma = 10^9$  e.s.u. and  $K = 3$ , the value of  $2\sigma f/K$  is only  $10^{-3}$  at  $f = 24,000$  Mc/s. so that the condition should be satisfied by any material likely to compose the moon's crust.

Consider now a semi-infinite, isotropic solid bounded by a plane infinite surface (Fig. 5). This may be used to represent a section of the lunar crust provided the area and depth concerned are small compared with the moon's radius. Take an element  $da$  of the boundary surface and consider the radiation

\* Intensity in a direction  $\theta$  is defined (following Milne,<sup>8</sup>) as the power per unit solid angle per unit projected area in the direction  $\theta$ .

emerging through  $da$  in directions confined within an elementary cone  $d\omega$ , whose axis is inclined at an angle  $\theta$  to the normal to the plane boundary. With the  $x$ -axis along the normal as shown, an elementary volume of the material is defined by the parallel planes  $x$  and  $x+dx$  and the cone. The radiation escaping through  $da$  is the sum of the emissions from all such elements, each weakened by absorption. Assuming scattering to be zero, the contribution from the element at  $x$  to the escaping radiation is

$$jdx(x \sec \theta)^2 \sec \theta d\omega d\zeta e^{-\int_0^x k \sec \theta dx} \dots \dots \dots (1)$$

where  $j$  is the radiation per unit volume per unit solid angle within the frequency limits relevant, and  $d\zeta$  is the solid angle subtended by  $da$  at the element.

Now  $d\zeta = \frac{da \cos \theta}{(x \sec \theta)^2}$  and taking  $k$  as a constant, the total intensity at  $da$  in the direction  $\theta$  is

$$I = d\omega da \int_0^\infty j e^{-\sec \theta kx} dx \dots \dots \dots (2)$$

If the material is everywhere in local thermodynamic equilibrium, Kirchhoff's law gives

$$j = \mu^2 k B(T),$$

where  $B(T)$  is the intensity of black-body radiation at the temperature  $T$  and in the frequency range considered. Since the Rayleigh-Jeans law holds in the microwave spectrum at the temperatures being considered, we may write  $\mu^2 B(T) = qT$ ,  $q$  being a constant. Thus if  $T(x)$  is the temperature at a distance  $x$  from the surface,  $j = kqT(x)$  for the elementary volume considered. Substituting in equation (2)

$$I = d\omega da k q \int_0^\infty T(x) e^{-\sec \theta kx} dx.$$

If the material were at a uniform temperature  $T_B$ , the intensity of the radiation would be

$$\begin{aligned} I_B &= d\omega da k q \int_0^\infty T_B e^{-\sec \theta kx} dx \\ &= \frac{d\omega da q T_B}{\sec \theta} \end{aligned}$$

which corresponds to radiation from a black body (Milne, S, p. 82). Thus the temperature of a black body emitting the same amount of radiation as the element of crust considered is

$$T = T_B \frac{I}{I_B} = k \sec \theta \int_0^\infty T(x) e^{-k \sec \theta x} dx \dots \dots \dots (3)$$

As the surface temperature  $T_s$  is a periodic function of time, of period  $2\pi/\Omega$ , it may be expressed as a Fourier series

$$T_s = T_0 + \sum_{n=1}^{\infty} T_n \cos n\Omega t.$$

Then the temperature at a depth  $x$  is given by (Carslaw and Jaeger, 6, p. 50)

$$T(x) = T_0 + \sum_{n=1}^{\infty} T_n e^{-\beta_n x} \cos (n\Omega t - \beta_n x) \dots \dots \dots (4)$$

where  $\beta_n = (n\Omega/2\alpha)^{\frac{1}{2}}$  is the thermal attenuation constant,  $\alpha = C/\rho s$  is the diffusivity and  $C$ ,  $\rho$ , and  $s$  are the thermal conductivity, density, and specific heat respectively.

This equation represents a series of temperature waves travelling into the crust, the harmonics diminishing in amplitude and shifting in phase more rapidly than the fundamental.

Consider the temperature distribution over any sphere concentric with the moon whose surface lies beneath that of the moon. The temperature-against-longitude curve will have a fundamental of period  $2\pi/\Omega$  and a series of harmonics. The integral over the lunar disk of any of these harmonics is always small and as our radio beam covers the whole disk, the efficiency of detection of harmonics is low compared with that of the fundamental. In addition, the harmonics are relatively small at the depth from which our radiation comes so that we may neglect all but the fundamental oscillation in the remainder of the analysis. The treatment with harmonics present is very similar.

Substituting for  $T(x)$  in equation (3) we have

$$T = k \sec \theta \int_0^\infty [T_0 + T_1 e^{-\beta x} \cos (\Omega t - \beta x)] e^{-k \sec \theta x} dx.$$

This may be rewritten

$$T = k \sec \theta \int_0^\infty T_0 e^{-k \sec \theta x} dx + \frac{1}{2} k \sec \theta \int_0^\infty T_1 e^{-\beta x - k \sec \theta x} \{e^{j(\Omega t - \beta x)} - e^{-j(\Omega t - \beta x)}\} dx$$

which after some reduction becomes

$$T = T_0 + \frac{T_1}{(1 + 2\delta + 2\delta^2)^{\frac{1}{2}}} \cos (\Omega t - \psi) \dots\dots\dots (5)$$

$$\text{where } \delta = \frac{\beta}{k \sec \theta} \text{ and } \tan \psi = \frac{\delta}{1 + \delta}.$$

Examination of equation (5) shows that the phase lag of the observed temperature has a maximum value of  $45^\circ$  to which it tends as the ratio of the attenuation of the thermal wave to that of the electromagnetic wave approaches infinity. At this limit, radiation is received from layers so far below the surface that the temperature is constant. However, the experimental value of the phase lag is  $45^\circ$  and the amplitude of oscillation is 0.39 that of the surface temperature variation. It is evident, therefore, that equation (5) does not provide a satisfactory picture of the actual conditions. It is likely that the results are explained by a type of lunar surface in which the ratio of electromagnetic to thermal absorption coefficients increases in some way with depth below the surface.

#### (d) Presence of a Layer of Dust

It would be difficult and probably not very fruitful to attempt an analysis similar to that resulting in equation (5) but with  $k/\beta$  a general function of depth. However, a case capable of simple treatment is one in which solid, uniform material is covered with a thin layer of poorly-conducting dust. Such a picture was suggested by Jaeger and Harper,\* who point out that at the low atmospheric

\* Personal communication 1948.

pressure existing on the moon any layer of dust or gravel would have a very low conductivity. They further show that Pettit's eclipse results may be explained by a thin layer of poorly conducting material. Earlier workers who have measured the polarization of light reflected from the moon's surface have suggested the possibility of a layer of dust (for example, Wright, 9).

Let the surface temperature of a semi-infinite solid, bounded by a plane and covered by a layer of dust, be given by

$$T'_s = T_0 + T_1 \cos \Omega t.$$

The temperature at a depth  $x$  is then

$$T(x) = T_0 + T_1 e^{-\beta x} \cos (\Omega t - \beta x).$$

The flux of heat across the  $x$  plane is given by

$$\begin{aligned} F &= -C \frac{\partial T(x)}{\partial x} \\ &= -C\beta T_1 e^{-\beta x} \{\sin (\Omega t - \beta x) - \cos (\Omega t - \beta x)\}. \end{aligned}$$

When  $x=0$  this reduces to

$$F = \sqrt{2} C \beta T_1 \cos (\Omega t + \pi/4).$$

If there is a very thin skin of material of thickness  $\tau$  and thermal conductivity  $C'$  covering the solid, the temperature drop through this skin will be

$$\frac{\tau}{C'} F = -\frac{\sqrt{2} \tau C \beta T_1}{C'} \cos (\Omega t + \pi/4).$$

The temperature at the external surface of the skin is therefore

$$T_s = T_0 + T_1 \left[ \cos \Omega t + \frac{\sqrt{2} \tau C \beta}{C'} \cos (\Omega t + \pi/4) \right]$$

which after simplification becomes

$$T_s = T_0 + T_1 (1 + 2\delta_s + 2\delta_s^2)^{\frac{1}{2}} \cos (\Omega t + \psi_s)$$

where  $\delta_s = \frac{\tau C \beta}{C'}$

and  $\tan \psi_s = \frac{\delta_s}{1 + \delta_s}$ .

If we assume that the absorption of microwaves by the skin is negligible, this temperature is to be compared in amplitude and phase with the "microwave temperature" of the solid which as before is given by equation (5)

$$T = T_0 + \frac{T_1}{(1 + 2\delta + 2\delta^2)^{\frac{1}{2}}} \cos (\Omega t - \psi).$$

The ratio of the amplitude of the temperature fluctuation is

$$(1 + 2\delta + 2\delta^2)^{\frac{1}{2}} (1 + 2\delta_s + 2\delta_s^2)^{\frac{1}{2}}$$

and the total phase lag of the microwave temperature behind the surface temperature is  $\psi_s + \psi$ .

It may be seen that for a given reduction of amplitude of the temperature oscillation, the maximum phase lag is given when  $\delta = \delta_s$  and  $\psi = \psi_s$ . For a total phase lag of  $45^\circ$ , as obtained experimentally, the amplitude reduction when  $\delta = \delta_s$  is about 3.4. This is to be compared with 2.6 obtained experimentally and suggests that the skin and the solid contribute roughly equal amounts to the phase lag and amplitude reduction.



If, then, we assume equal effects due to the surface skin and the solid layer below, the values of the constants defining amplitude and phase changes are given by

$$\begin{aligned}\psi &= \psi_s = 22.5^\circ \\ \delta &= \delta_s = 0.71.\end{aligned}$$

Substituting  $\delta_s = \frac{\tau C \beta}{C'} = \frac{\tau}{C'} \left( \frac{\Omega C \rho s}{2} \right)^{\frac{1}{2}}$  and putting  $\Omega = 2.66 \times 10^{-6}$  radian sec.<sup>-1</sup>

we find  $\frac{\tau}{C'} (C \rho s)^{\frac{1}{2}} \sim 600$ .

The density and specific heat of the solid material of the moon's crust will probably not differ essentially from those for basalt. The conductivity, however, may have any value between that for solid basalt ( $4.2 \times 10^{-3}$  cal. cm.<sup>-2</sup> sec.<sup>-1</sup>) and a very low value, say  $10^{-5}$  or  $10^{-6}$ . The latter values might be expected if the lunar crust is broken up, the conductivity then being greatly reduced in the absence of an atmosphere. In Table 1, values of  $\tau$ , the thickness of the dust layer, are given for various values of  $C$  and hence of  $(C \rho s)^{-\frac{1}{2}}$ .

TABLE 1

$C$	$(C \rho s)^{-\frac{1}{2}}$	$\tau/C'$	$C'$	$\tau$ (cm.)
$4.2 \times 10^{-2}$	20	$1.2 \times 10^4$	$10^{-6}$	0.012
"	"	"	$10^{-5}$	0.12
$2.0 \times 10^{-4}$	92	$5.5 \times 10^4$	$10^{-6}$	0.05
"	"	"	$10^{-5}$	0.55
$1.0 \times 10^{-5}$	410	$2.5 \times 10^5$	$10^{-6}$	0.25

It is evident that if the material of the moon's crust is solid, with a conductivity similar to that of terrestrial rocks or lava, then the dust film will be very thin, perhaps of the order of one millimetre.

The origin of such a film is explicable in terms of meteors falling on to the moon's surface. It has been indicated by Watson (10, p. 116) that a summation of all the masses of meteoric material falling on to the earth or its atmosphere provides about 1000 kilograms per day, which in a period of  $2 \times 10^9$  years would lead to the formation of a layer less than one centimetre thick. Less material per square centimetre of surface should fall on to the moon because of its smaller gravitational pull but it might also be necessary to consider the fragments broken from the lunar crust at impact of meteors and meteorites.

It may be of interest to estimate the average depth below the moon's surface from which our microwave radiation emerges. If we take  $\psi = \psi_s = 22\frac{1}{2}^\circ$  as above, then  $\delta = \delta_s = 0.71$ . The solid material may have a thermal diffusivity not far from that of lava ( $0.004$  cm.<sup>2</sup> sec.<sup>-1</sup>) and putting  $\Omega = 2.66 \times 10^{-6}$  radian sec.<sup>-1</sup> we have  $\beta = 0.018$ , so that  $k \sec \theta = \beta/\delta = 0.025$ .

Using Snell's law and assuming a likely value of  $K=5$ , it may be shown that the average value of  $\sec \theta$  over the disk is about 1.05. We have, therefore, the absorption coefficient  $k=0.025$ . Using Milne's concept of average optical depth  $x$  (8, p. 84), from which radiation may be considered to originate, we have  $kx=1$ . This results in a value of  $x=40$  centimetres.

Again, since  $k = \frac{4\pi\sigma}{cK^4}$ , we find the corresponding value of electrical conductivity  $\sigma=1.3 \times 10^8$  e.s.u.; a figure which corresponds to the conductivity of average soil.

## VI. CONCLUSIONS

We find that microwave thermal radiation from the moon originates in layers below the surface, differing in that respect from infra-red radiation. Because of this, the phase angle of the lunar microwave temperature curve lags by about  $45^\circ$  behind the curves of surface temperature and surface radiation. The results are consistent with the existence of a layer of dust covering the solid material of the moon's crust, the average depth from which radiation emerges being about 40 centimetres.

Interpretation of the microwave results, in terms of thermal waves propagating into the interior of the moon, leads to values of temperature in the deep interior of  $234^\circ\text{K.}$  and average surface temperature of the disk of the new moon of  $145^\circ\text{K.}$  with an accuracy better than  $\pm 10$  per cent.

## VII. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Radiophysics, C.S.I.R. Considerable assistance, both in the development of the equipment and in the observational work, was given by Mr. J. V. Hindman.

## VIII. REFERENCES

- (1) DICKE, R. H.—*Rev. Sci. Instrum.* **17**: 268 (1946).
- (2) DICKE, R. H., and BERINGER, R.—*Astro. J.* **103**: 375-6 (1946).
- (3) PETTIT, E., and NICHOLSON, S. B.—*Ibid.* **71**: 102-35 (1930).
- (4) PETTIT, E.—*Ibid.* **81**: 17-36 (1935).
- (5) PETTIT, E.—*Ibid.* **91**: 408-20 (1940).
- (6) CARSLAW, H. S., and JAEGER, J. C.—"Conduction of Heat in Solids." 1st Ed. (Oxford, 1947.)
- (7) DICKE, R. H., BERINGER, R., KYLE, R. L., and VANE, A. R.—*Phys. Rev.* **70**: 340-8 (1946).
- (8) MILNE, E. A. "Thermodynamics of the Stars." *Handb. Astrophysik* **5**: 3 (1930).
- (9) WRIGHT, F. E.—*Proc. Nat. Acad. Sci.* **13**: 535-40 (1927).
- (10) WATSON, F. G.—"Between the Planets." 1st Ed. (The Blakiston Co.: Philadelphia, 1945.)

## APPENDIX I

### *Conversion to Lunar Temperatures*

As explained in Section III (c), the calculation of the true average temperature over the lunar disk requires a knowledge of both aerial properties and the distribution of microwave radiation. It has been found, however,

from a complete analysis which is too lengthy to include here, that the exact form of the law relating microwave radiation with lunar latitude  $\varphi$  and longitude  $\lambda$  relative to the subsolar point does not greatly affect the observed aerial temperatures. From Pettit and Nicholson's results the temperature of an element of the moon's surface while illuminated by the sun can be reasonably represented by  $T_s \cos^{\frac{1}{2}} \rho$ , where  $\rho$  is the angle between the normal to the element and the direction of the sun and  $T_s$  has a value of 374 °K. If for simplicity it is assumed that the moon's equator and its orbit about the earth are both in the plane of the ecliptic, this expression may be expanded as  $T_s \cos^{\frac{1}{2}} \varphi \cos^{\frac{1}{2}} \lambda$ .

The temperature of a point in terms of the relative longitude will be a periodic function represented by  $\cos^{\frac{1}{2}} \lambda$  while the point is illuminated and by some unknown cooling law during lunar night. The temperature of the point shortly before lunar sunset is proportional to  $\cos^{\frac{1}{2}} \varphi$  and subsequent cooling is slow, as shown by an analysis of Pettit and Nicholson's eclipse results. We assume, therefore, that the time-average (throughout a lunar cycle) of the temperature of a point of latitude  $\varphi$  on the moon's surface is proportional to  $\cos^{\frac{1}{2}} \varphi$ . The time-average of temperature of a point below the surface will be equal to that of the surface point above it.

The directional properties of the aerial were measured by two methods which will be described elsewhere. One of them depended on the comparison of the amounts of energy received from the sun's disk by the aerial under consideration and a broad-beam aerial of known gain. The sun's disk was assumed to be a uniform radiator.

These aerial measurements together with the law of temperature distribution given above allowed calculation of the space-average temperature over the lunar disk and also of the temperature of the point on the equator nearest the earth.

### *Lunar Emissivity*

The temperatures thus obtained are those which would provide the observed radiation if the moon were a black body. At the frequency used, however, the moon is probably not a black body; that is, if radiation were incident normally upon it, a proportion  $R$  of the total incident power would be reflected. The value of  $R$  is given, for poor conductors, by

$$R = \left( \frac{K^{\frac{1}{2}} - 1}{K^{\frac{1}{2}} + 1} \right)^2$$

where  $K$  is the dielectric constant, and may safely be taken as about 5, so that  $R$  is 0.147.

At angles of incidence other than normal, the reflection coefficient has different values which depend on the polarization of the energy. An examination shows, however, that it is a reasonable approximation to take the average value of  $R$  over the lunar disk as equal to the value in the centre. In order to allow for a probable gradient of density from the moon's surface into the interior, a slightly reduced value of  $R$  is taken, namely, 0.1. All lunar temperatures have, therefore, been increased above black-body values to allow for an emissivity of 0.9.

# EXPERIMENTS IN SEEDING CUMULIFORM CLOUD LAYERS WITH DRY ICE

By E. J. SMITH\*

(Plates 1-14)

[Manuscript received October 21, 1948]

## Summary

This paper describes five experiments in which granulated dry ice was dropped into layers of cumulus cloud at temperatures below freezing. In each case precipitation of snow or ice crystals occurred, and part or all of the cloud dispersed. Displays of haloes and similar phenomena were observed.

## I. INTRODUCTION

The mechanism by which rain is formed in a supercooled water cloud has been described by Bergeron(1). According to his theory the initial release of precipitation occurs when ice crystals form in or are introduced into the cloud.



Fig. 1.—Area in which experiments were carried out.

Owing to the difference in vapour pressure over water and ice, water evaporates from the supercooled drops and sublimates on to the ice crystals, which grow and fall, melting as they pass the freezing level.

\* Division of Radiophysics, C.S.I.R.



TABLE I

Date	27.iv.48	11.vi.48	17.vi.48	25.viii.48	25.viii.48
Latitude .. ..	35° 00'	32° 45'	32° 50'	33° 15'	33° 20'
Longitude .. ..	150° 30'	150° 50'	148° 40'	149° 30'	149° 30'
<i>Cloud—</i>					
Height of top (ft.) ..	11,500	13,500	12,200	8,200	8,200
Temperature of top (°C.) .. ..	—7	—10	—6½	—7	—7
Height of bottom (ft.) ..	9,500	12,000	11,200	4,500	4,500
Temperature of bottom (°C.) ..	—2	—9	—6	—3	—3
Lapse rate at top (°C./100 M.) ..	—0.2	0.3	0.5	0.5	0.5
Density .. ..	Dense	Moderate	Light	Dense	Dense
<i>Dry Ice Drop—</i>					
Time of release (hr.) ..	1501	1341	1044	1006	1003
Weight dropped (lb.) ..	60	40	25	20	40
Rate of release (lb./min.) .. ..	60	25	25	40	40
Pattern .. ..	Cross	Cross	Line	Line	Line
<i>Precipitation—</i>					
Time to start (min.) ..	16	12	9	10	9
Form .. ..	Snow, ice crystals, and rain	Ice crystals	Ice crystals	Snow	Snow
Depth of fall (ft.) ..	5,000	2,000	1,000	To ground	To ground
Approx. rate of fall (ft./min.) .. ..	800	500	300	800	600
Duration (min.) ..	18	18 or more	10	22	50
<i>Change in Cloud from Water to Ice—</i>					
Time to start (min.) ..	16	12	9	20	Unknown
Amount of cloud changed ..	All	Half	One-hundredth	4 sq. miles	50 sq. miles
Duration of change (min.) .. ..	6	18	10	10	50
Had change stopped spreading? ..	Yes	Unknown	Yes	Yes	Yes
<i>Optical Effects Observed—</i>					
Time (hr.) .. ..	1554	1407	1104		
22° halo .. ..	White	Coloured	Coloured		
46° halo .. ..	White	—	—		
25° parhelia .. ..	—	Coloured	Coloured		
90° parhelia .. ..	—	—	Faint, white		
Parhelic circle .. ..	—	Faint, white	White	None	None
Upper arc of contact ..	White	—	—		
Lower arc of contact ..	White, brilliant	—	—		
Specular reflection below sun ..	Possibly	Faint, white	Faint, white		
Elevation of sun ..	15° 30'	26° 28'	31° 46'		

In many experiments carried out in different parts of the world ice crystals have been artificially introduced into supercooled water clouds by seeding with dry ice. Descriptions of such experiments have been given by Schaefer(2), who seeded stratus clouds, and Kraus and Squires(3), who seeded towering cumulus. Both experiments produced artificial precipitation: in the latter, the cloud rapidly grew and developed into a cumulonimbus.

A comprehensive programme of investigation of artificial rain formation by the dry ice process is being carried out in this Laboratory. Some results of this investigation was published in *Nature*(3) and a more complete description is in preparation. The present paper describes five occasions on which artificial precipitation was induced to fall. They are described because they give striking evidence of the way in which ice crystals and snow-flakes form in a cloud of supercooled water droplets when it is seeded with dry ice. A summary of the results is given in Table 1.

## II. GENERAL DESCRIPTION OF EXPERIMENTS

The experiments were carried out over New South Wales on April 27, June 11, June 17, and August 25, 1948. A map of the area in which they were performed appears in Figure 1. The dry ice was dropped and the results observed from a R.A.A.F. Dakota based at Richmond, 30 miles north-west of Sydney.

Before each experiment was performed temperature soundings of the atmosphere were taken in the aircraft during the initial ascent. They were made with a resistance thermometer accurate to  $1^{\circ}\text{C}$ . No attempt was made to measure cloud temperature and all measurements quoted in this paper were made in clear air. When a suitable cloud had been selected, careful observation was made to be sure that neither it nor surrounding clouds were precipitating or likely to produce precipitation naturally. When this point had been ascertained, the aircraft flew just above the top of the selected cloud and the dry ice was dropped. The aircraft then remained in the neighbourhood, usually just below the cloud base, to observe results.

## III. EXPERIMENT ON APRIL 27, 1948 (Figs. 2-4; Plates 1-7)

(i) *General Meteorological Situation.*—A diffuse warm front lay along the New South Wales ranges, roughly parallel to the coast. Scattered light rain was falling in the coastal region but not in the immediate vicinity of the experiment. The winds aloft were weak and variable, with a general easterly stream in the first five thousand feet, tending north east and then north-west above. The sounding in Figure 2, taken in the aircraft during the climb south from Sydney to the site of the experiments, shows inversions at 6500 and 12,000 ft.

(ii) *Description of Cloud Seeded.*—A layer of altocumulus extended inland for some distance from the coast. One part of this, ten or fifteen miles inland, was selected for the experiment. In this region the layer was broken and amounted to about six-tenths at the beginning of the experiment and less at the end. The selected cloud was 2000 ft. thick and about three miles long by one and a half miles broad, with its major axis east and west. It is shown in Plates 1 and 2. The top was at 11,500 ft. ( $-7^{\circ}\text{C}$ ., 657 mb.) and the base at 9500 ft.

( $-2^{\circ}\text{C.}$ , 710 mb.) (see Fig. 2). The layer was therefore below the upper inversion and presumably arose by convection in the layer between the inversions. It was decaying at the time of the experiment, for the whole layer was dispersing and the base was ragged. It was a dense, opaque cloud, exhibiting no trace of ice. Except at the edges, the structure appeared fairly uniform. The top showed undulations but not to the extent of separate cells. The outlines were perfectly clear. Although ragged, the bottom showed no signs of precipitation.

There were other clouds in the neighbourhood: scattered low cumulus and high altocumulus or cirrocumulus, a layer of which extended northwards from the area of the experiment. Out to sea, thirty or more miles away, were large cumulus and cumulonimbus, many of them raining.

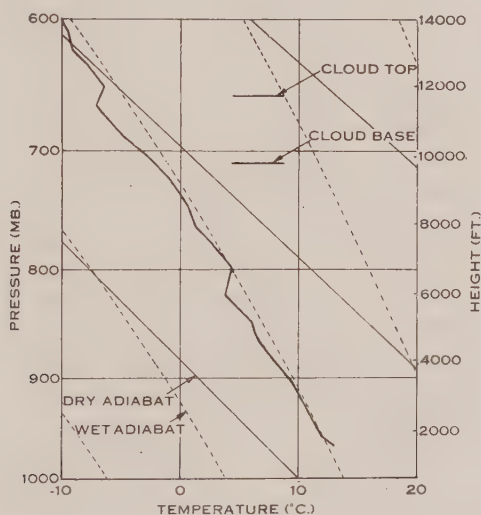


Fig. 2.—First experiment, April 27, 1948.  
Aircraft sounding taken in clear air during the climb.

(iii) *Dry Ice Drop*.—Sixty pounds of granulated dry ice were dropped into the selected cloud at 1501 hours. Two runs were made at about 500 ft. above the top of the cloud and dry ice was dropped, at the rate of 60 pounds per minute, on two lines about  $1\frac{1}{2}$  miles long in the form of a cross, as shown in Figure 3.

(iv) *Precipitation Observed*.—The aircraft then descended to the bottom of the cloud and awaited results. Nothing was seen for 16 minutes, after which vertical streaks of snow (shown in Plate 3) were observed to fall from the western end of the cloud, which rapidly changed its appearance from a clearly defined waterdrop cloud to an amorphous mass of ice crystals. The change spread throughout the cloud until 22 minutes after the dry ice was dropped, by which time the whole cloud appeared to consist of ice crystals, with snow falling from the bottom as shown in Plate 4. Thirty-four minutes after the dry ice was dropped, the snow had ceased falling and the cloud was a semi-transparent mass

of ice crystals, covering roughly its original area and extending from about 11,000 to 9000 ft.

The aircraft flew through the snowstorm at its peak, and the snow-flakes could be clearly seen. Their maximum diameter was estimated as about one centimetre. After the snow stopped falling the aircraft flew through the cloud several times. The ice crystals could be seen glinting in the sun, but appeared to be very small.

The snow fell to about 1090 ft., changing to rain on the way, but appeared to evaporate before reaching the ground. The cloud of ice crystals gradually became less dense until 49 minutes after the dry ice drop, when it had almost disappeared. It is shown dissipating in Plates 5 and 6.

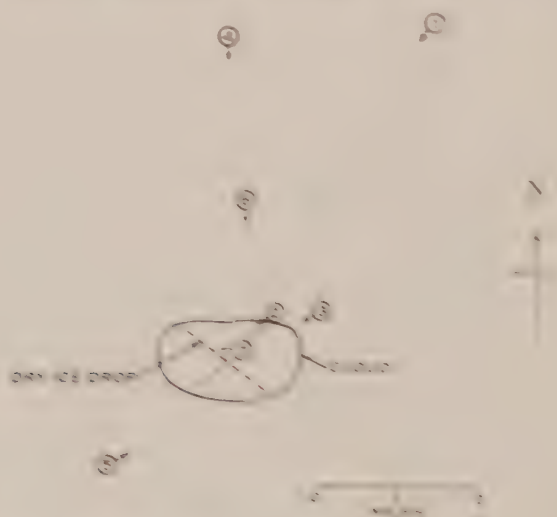


Fig. 3.—Key to Plates 1-7. First experiment.  
Plan positions, relative to cloud, from which  
photographs were taken.

(1) *Optical Phenomena*.—As the aircraft flew through the cloud of ice crystals, a display of optical phenomena was observed around the sun, as shown in Plate 7 and Figure 4. The pattern appeared and disappeared at the position of the aircraft changed, and it was sometimes only partly visible. The complete 22° halo, two bright arcs of contact, and the 44° halo, none or but sometimes could be seen. They were white, no colour being observed. There were no signs of parhelia. These phenomena and their causes are described by Humphreys(4).

Directly below the sun there was a brilliant white arc of contact(4, p. 426), which is shown in Plate 7. Its brilliance was comparable with that of the sun itself and was greater than that of the haloes. It took approximately the form of an ellipse, the upper portion of which was extremely bright and the lower portion less so. The upper portion was at 22° from the sun, making the 22° halo, which was clearly visible to the eye but can only just be seen in the photo-



graph. At the bottom of this "ellipse" there was a fainter, but still quite bright spot. This may have been the point of intersection of the two arms of the arc of contact or a specular reflection of the sun in horizontal crystal faces. In the centre of the photograph are some faint octagons caused by the iris diaphragm of the camera.

#### IV. EXPERIMENT ON JUNE 11, 1948 (Fig. 5; Plate 8)

(i) *General Meteorological Situation.*—On the afternoon of June 11, 1948, a col lay over the central coast of New South Wales. In the area of the experiment, to the north of Sydney, the winds aloft were west-north-west to north-west, increasing slowly with height. The radiosonde sounding taken at Rathmines at 2100 hours (seven hours later) agreed in general characteristics with that

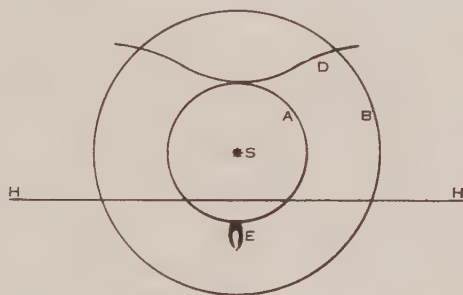


Fig. 4.—Halo display during first experiment.  
S, Sun. Elevation  $15^{\circ} 30'$ ; H,H, Horizon;  
A,  $22^{\circ}$  Halo; B,  $46^{\circ}$  Halo; D, Upper arc of  
contact; E, Lower arc of contact.

taken in the aircraft during the climb northwards from Richmond. Both showed a lapse rate very nearly equal to the wet adiabatic from 5000 to 15,000 ft. The radiosonde humidity curve indicated that the air mass was almost saturated from 10,000 to 15,000 ft.

(ii) *Description of Cloud Seeded.*—Eighty miles north of Sydney there was an isolated layer of altocumulus about ten miles long and three miles wide, aligned north and south. Its widest and thickest part was towards the northern end, where the height and temperature of the top were 13,500 ft. and  $-10^{\circ}\text{C}.$ ; the bottom was at 12,000 ft.,  $-9^{\circ}\text{C}.$  Freezing level was at 8000 ft.

The cloud was moderately dense and composed of cells each a quarter to half a mile across. It was just possible to see the ground between the cells, which themselves were opaque. The cloud showed no trace of ice and the bottom was flat with no sign of precipitation.

There were a few traces of low cumulus in the same neighbourhood, and other layers of cloud at distances of twenty miles or more. About forty miles to the west there was a large cumulus which developed into a cumulonimbus during the test, and at greater distances in all directions were other cumulus and cumulonimbus formations.

(iii) *Dry Ice Drop*.—Forty pounds of granulated dry ice were dropped into this cloud at 1341 hours from a height of 14,000 ft. Two runs were made over the thickest part of the cloud and the dry ice was dropped on lines two and a half miles long in the form of a cross.

(iv) *Precipitation Observed*.—Twelve minutes later precipitation streaks were seen to begin falling from the cloud and a change was propagated through it. One of the cloud cells suddenly changed its appearance from water drops to ice crystals; ice fell out of the bottom, falling some 2000 ft. before evaporating, and the cell itself became thin and transparent. This whole process took about six minutes. While it was taking place a neighbouring cell began the same cycle. Thus the change from water to ice with precipitation spread through the cloud, not steadily, but in steps as one cell after another suddenly changed. This process continued until about half an hour after the dry ice was dropped.

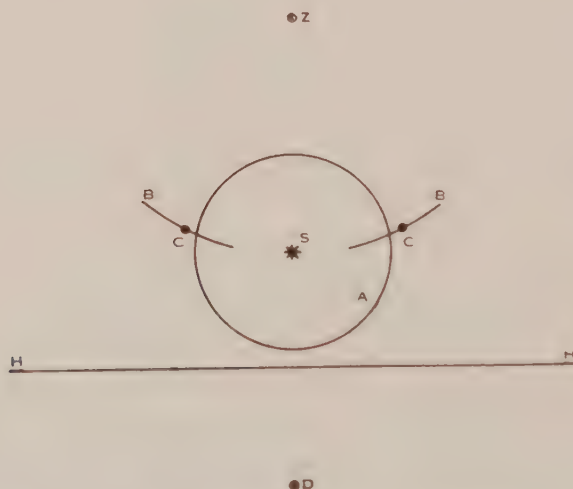


Fig. 5.—Halo display during second experiment.  
 S, Sun. Elevation  $26^{\circ} 28'$ ; H,H, Horizon; Z, Zenith;  
 A,  $22^{\circ}$  Halo (coloured, bright); B,B, Area of parhelic circle  
 (white, rather faint); C,C, Parhelia at  $24^{\circ}$  (coloured, bright);  
 D, Specular reflection (white, faint).

During this time the whole of the infected area, some three miles by four, changed into a transparent and wispy mass of ice crystals, with streaks of precipitation still visible in a few places. The southern end of the cloud and a few scraps on the northern end were unchanged. They may have been affected later but observation had to cease at this stage.

Twenty-four minutes after the drop the aircraft flew through the precipitation, which appeared to consist of quite small ice crystals with no snow-flakes.

(v) *Optical Phenomena*.—A display of optical phenomena, shown in Plate 8 and Figure 5, was observed. The  $22^{\circ}$  halo was bright, complete, and was brilliantly coloured. There were no other haloes. Two parhelia were seen,

also brightly coloured, level with the sun, about  $24^\circ$  from it. A faint trace of the parhelic circle could be seen in the neighbourhood of the parhelia. No arcs of contact were seen. A faint specular reflection of the sun vertically below it was seen momentarily and took the form of a clear image which was below the horizon by an angle equal to the sun's elevation.

#### V. EXPERIMENT ON JUNE 17, 1948 (Fig. 6; Plate 9)

(i) *General Meteorological Situation.*—A ridge of high pressure lay across eastern Victoria and western New South Wales, aligned north-west to south-east. A moderately deep cyclone was centred across the north Tasman Sea, causing a strong, deep south-east stream to blow on to the New South Wales coast, where there was confused cumulus cloud, and light to moderate rain was falling. Inland the air became warmer, the coastal cloud mass broke up, and there was some light rain on the coastal slope of the ranges.

(ii) *Description of Cloud Seeded.*—There was a layer of altocumulus cloud 100 miles west of Sydney, some forty miles long and three miles wide, aligned east and west. It could be described as a string of small clouds, of which the largest and densest was about fifteen miles by four. The top was at 12,200 ft.,  $-6\frac{1}{2}^\circ\text{C.}$ , and the bottom at 11,200 ft.,  $-6^\circ\text{C.}$  Freezing level was at 8000 ft. The lapse rate near the cloud was very nearly the wet adiabatic from 8000 ft. to 200 ft. above the cloud top, at which level measurements ceased. The widest and thickest part was situated one-third of the length from the western end, and both ends tapered off to nothing in depth and width. It was a thin cloud; the ground was just visible through its densest parts. It was composed of patches or cells about a quarter-mile wide which were only just joined together, i.e. between the cells it was easily possible to see the ground through the cloud, while some cells were not joined at all. There was no indication of the presence of ice crystals. The bottom was flat with no sign of precipitation.

At about 5000 ft. there was a layer of two-tenths stratocumulus arranged in orderly rows which became denser and higher to the east, while to the north about fifty miles away there was a line of towering cumulus.

(iii) *Dry Ice Drop.*—At 1044 hours, 25 pounds of granulated dry ice were dropped into the layer from 12,500 ft., on a line three miles long from north to south across the thickest part of the cloud. The aircraft then flew level with the bottom of the cloud.

(iv) *Precipitation Observed.*—Nine minutes after the drop, a small streak of precipitation was seen as ice crystals fell out of the bottom of the cloud. This precipitation grew, and 15 minutes after the dry ice drop it covered about half a square mile, confined to the northern end of the line along which the dry ice was dropped. The precipitation came from a few of the cells of the main cloud and the disturbance did not spread to the rest. The cells affected turned rapidly to an amorphous mass of ice crystals as precipitation took place, and then remained as a transparent ice-crystal cloud after it had stopped. The precipitation appeared to take the form of small ice crystals. The aircraft flew through it several times and no snow-flakes were observed. The precipitation fell about 1000 ft. below the bottom of the cloud before evaporating.

About 45 minutes after the drop, the part of the cloud which had precipitated had almost dispersed, being visible from a short distance as a thin haze of ice crystals. Immediately below it there was at this time another cloud of ice crystals. Apparently some of the precipitation fell about 1000 ft. and remained there for one hour after the drop in the form of a transparent mass of ice crystals similar in appearance to, but separated from, the one above. No sign of the disturbance spread to the rest of the cloud, which was quite unchanged.

(v) *Optical Phenomena*.—As the aircraft flew through and past the precipitation and the clouds of ice crystals, displays of haloes and associated phenomena were observed, as shown in Plate 9 and Figure 6. There was a bright and strongly coloured  $22^\circ$  halo, complete. No other halo was seen nor

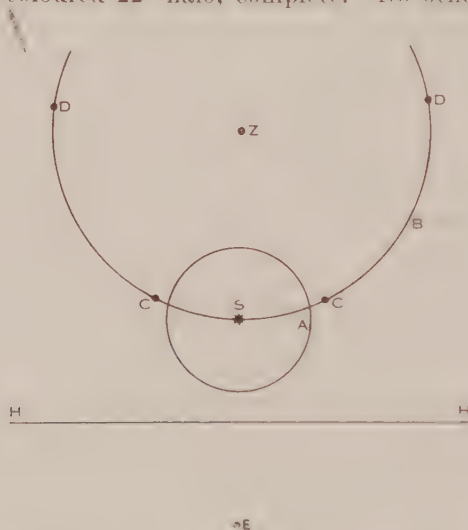


Fig. 6.—Halo display during third experiment.  
S, Sun. Elevation  $31^\circ 46'$ ; H,H, Horizon;  
Z, Zenith; A,  $22^\circ$  halo; B, Parhelic circle;  
C, Parhelia at  $25^\circ$ ; D, Parhelia, position not  
measured, perhaps at  $90^\circ$ ; E, Specular reflection.

was any arc of contact. There was a white parhelic circle of varying extent, sometimes about two-thirds complete. Two brightly coloured parhelia appeared at about  $25^\circ$  and two more, faint and white, at about  $90^\circ$ . There was a faint specular reflection below the sun in the form of a fairly clear image, which was below the horizon by an angle equal to the sun's elevation.

## VI. EXPERIMENTS ON AUGUST 25, 1948

(i) *General Meteorological Situation*.—A decaying cyclone was centred near Tasmania and a moderate westerly wind was blowing over New South Wales. A layer of eight-tenths cumulus cloud covered the mountain ranges inland from the coast of N.S.W., the cloud tops being at about 8000 ft. and their base at 4500 ft. Two very light local falls of natural snow were observed, but none occurred within thirty miles of the area of the experiments. The tops of the



clouds showed the form characteristic of water drops, except those from which snow was falling, which had the "fuzzy" appearance of ice clouds. Much heavier snow had fallen the day before and was still lying on high ground. Figure 7 shows a sounding taken in the experimental area, where mountains range from 2000 to 4000 ft.

(ii) *Description of Clouds Seeded.*—The two cloud peaks selected for experiment were about fifteen miles north-west of Bathurst and about fifty miles west of the edge of the cloud layer, which became thinner and less dense over the easternmost twenty miles. They were about ten miles apart and appeared exactly similar to all the other peaks. Each was about two miles across at the top and five miles across at the base, and in places they were joined to their neighbours. There was no higher cloud. No trace of icy cloud was

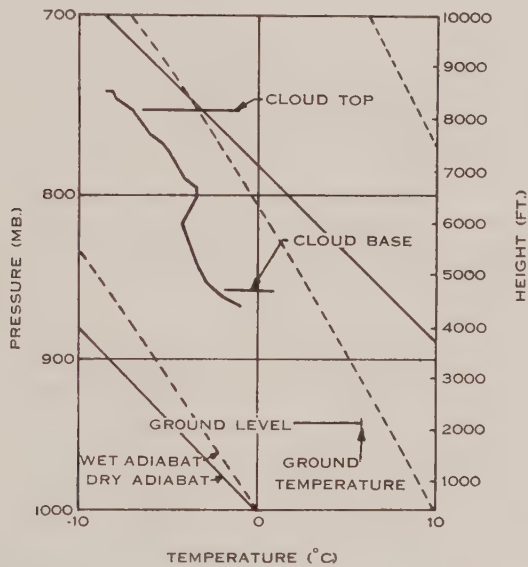


Fig. 7.—Experiments on August 25, 1948. Aircraft sounding taken in Bathurst area.

visible for at least thirty miles and no precipitation of any kind could be seen from under the clouds, where visibility was excellent. Each peak was the subject of a separate experiment. Photographs taken as the aircraft ran up to drop dry ice appear in Plates 10 and 12.

(iii) *Fourth Experiment, Dry Ice Drop.*—Twenty pounds of dry ice were dropped at 1006 hours during a straight run of one and a half miles just above the top of the cloud. The aircraft's height was 8300 ft., the temperature  $-7^{\circ}\text{C}.$ , and the course north-east.

(iv) *Precipitation Observed from Fourth Cloud.*—Small streaks of precipitation appeared after 10 minutes. At this time the aircraft was within a few hundred yards of them, and they would not have been visible for very far. Twelve minutes after the dry ice drop, the precipitation had become much more con-

spicuous; it extended about 500 ft. below the cloud base and would have been visible for perhaps 10 miles. After 18 minutes the precipitation reached the ground 15 miles north of Bathurst. It was at its heaviest after 20 minutes, when it covered about four square miles. The aircraft flew through the precipitation, 500 ft. below the cloud base, and it was seen to consist of densely-packed snow-flakes estimated visually to be 2-5 mm. in diameter. The snow-storm, shown in Plate 11, was opaque. There was a fairly well-defined level at 3500 ft. where the precipitation became less opaque, probably because of the melting of the snow-flakes as they passed the freezing level. Some of the snow fell over mountains and appeared to reach the higher peaks without melting. After 24 minutes the precipitation became less intense. The aircraft flew through it, and the snow-flakes, though still closely packed, were now about 1 mm. in diameter. After 32 minutes the precipitation had stopped.

(v) *Dispersal of Fourth Cloud.*—The cloud began to disperse after 20 minutes when the precipitation was at its height, and it had almost disappeared after about 25 minutes. The appearance then was remarkable, as quite heavy precipitation seemed to be falling from a gap in the cloud. The discrepancy in time between the disappearance of the cloud and the stopping of the precipitation is, of course, due to the time the precipitation took to fall.

(vi) *Fifth Experiment, Dry Ice Drop.*—Forty pounds of dry ice were dropped at 1003 hours in a straight run two and a half miles long on a north-easterly course. The aircraft flew just above the top of the cloud at 8300 ft. where the temperature was  $-7^{\circ}\text{C}$ .

(vii) *Precipitation Observed from Fifth Cloud.*—After nine minutes very small streaks, visible for about one mile, fell from the bottom of the cloud. Sixteen minutes after the dry ice drop, the streaks extended 1000 ft. below the cloud and were visible for about 10 miles. After 19 minutes the snow-storm shown in Plate 13 had developed and reached the ground about 5 miles north of Bathurst. It continued to grow in intensity and area, covering approximately one square mile after 20 minutes, four square miles after 29 minutes, 15 square miles after 35 minutes, and 50 square miles after 50 minutes, when the edge extended to about the area of the fourth experiment. The intensity grew to a maximum after some 40 minutes, when the clouds drifted from a plain about 2000 ft. above sea-level to 4000-ft. mountains. The precipitation, shown in Plate 14, again took the form of snow which appeared to melt before reaching the ground. As the cloud drifted above the mountains the precipitation decreased in intensity but the snow appeared to reach the ground without melting, though it could not be seen lying there because of snow which had fallen the day before. Fifty-two minutes after the dry ice was dropped the whole cloud was over the mountain and the precipitation was quite light. It still covered about fifty square miles but it was just possible to see through it. The last traces of precipitation were seen after about an hour.

The aircraft flew through the snow 29 minutes after the dry ice drop. The flakes were very closely packed and about 2 mm. in diameter. On the edge of the main shower some snow-flakes about 5 mm. in diameter and much less

closely spaced were encountered. A second flight through the snow was made after 43 minutes, when the flakes were about 1 mm. in diameter and still very closely packed.

(viii) *Dispersal of Fifth Cloud.*—It was not easy to observe the fifth cloud. The aircraft flew below the level of the cloud base in order to observe the precipitation, which as it spread obscured the cloud which had been precipitating for an appreciable time. Consequently the history of the cloud in the early stages is not known. When the aircraft flew through the snow 43 minutes after the dry ice drop, the centre of the cloud appeared to have practically dispersed. The aircraft flew at 4500 ft., just below the original level of the cloud base, and judging by the light intensity, there was not more than a few hundred feet of cloud or snow above the aircraft. When the snow became transparent, about 55 minutes after the drop, a hole could be seen in the cloud layer roughly corresponding with the snow. After 70 minutes, the snow having stopped, the aircraft climbed through this hole, which was then about seven miles across, so that the general situation could be observed from above. The clouds had by this time drifted so far east that they had reached the region where the whole cloud layer was beginning to disperse.

## VII. COMPARISON OF RESULTS

(i) *Changes in Cloud.*—In the first three experiments the seeding with dry ice caused a change from a cloud composed of water drops to one composed of ice crystals. In the first cloud, which was dense and continuous, the change spread rapidly throughout. In the second, which consisted of cells rather thinly joined together, the change spread more gradually, cell by cell. In the third, a thinner cloud with cells less well joined together, the change did not spread. Observations of changes in the clouds in the last two experiments could not be made owing to the surrounding layers.

In the first three cases, the parts of the cloud which changed to ice crystals soon dispersed to a fairly transparent state, taking five or ten minutes in the small cells and fifteen minutes in the more solid cloud of April 27. After this the clouds very gradually became more transparent, but they persisted in the form of small ice crystals as long as observations could be made. The ice crystals could be seen as faint wispy clouds at short range, and by the optical effects produced as the aircraft flew through them. Similar changes appear to have taken place in the clouds seeded on August 25, 1948, though here the observations are less conclusive.

(ii) *Precipitation.*—There appears to have been good qualitative agreement between the density of the cloud and the amount of precipitation. Thus on April 27 and August 25 solid clouds yielded fairly heavy snow-storms, on June 11 a much less dense cloud yielded a moderate shower of ice crystals, and on June 17 a thin cloud yielded a light shower of ice crystals. The distance through which the precipitation fell was also greater for the denser clouds.

The rate of fall of the first visible precipitation was not measured accurately for the first three cases but was estimated as 800, 500, and 300 ft. per minute



respectively. For the two experiments on August 25, the rate of fall was 800 and 600 ft. per minute. This probably indicates that larger crystals fell from the denser clouds.

(iii) *Optical Phenomena*.—The different displays of haloes etc. are interesting. According to Humphreys(4) they indicate that on April 27 there were columnar crystals floating with their axes horizontal, while on June 11 and 17 the crystals, which caused the effects seen, were probably mainly tabular with vertical axes.

Photographs show that the parhelia were outside the  $22^\circ$  halo by about the theoretical angle as given by Humphreys(4, p. 513). The theoretical displacements were  $2^\circ$  for June 11 and  $3^\circ 10'$  for June 17. The corresponding angle for April 27 is about  $0^\circ 45'$ , but no trace was seen of parhelia even as a local brightening of the halo.

### VIII. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Radiophysics, C.S.I.R. Grateful acknowledgment is made to the Royal Australian Air Force for the use of aircraft in the experiments and to the R.A.A.F. pilots who flew them. The general meteorological data were obtained by Mr. P. Squires of this Division, who also gave much valuable advice.

### IX. REFERENCES

- (1) BERGERON, T.—*P.V. Mët. Un. Géod. Géophys. Int.* Lisbon (1933).
- (2) SCHAEFER, V. J.—Gen. Elec. Res. Lab. Quart. Prog. Rep. No. 1. July, 1947.
- (3) KRAUS, E. B., and SQUIRES, P.—*Nature* **159**: 489 (1947).
- (4) HUMPHREYS, W. J.—“Physics of the Air”, pp. 501-54. (McGraw-Hill: New York, 1940.)

### EXPLANATION OF PLATES 1-14

#### PLATE 1

First experiment, April 27, 1948, 1455 hours. Looking SW. at 11,700 ft. The cloud in which dry ice was dropped, and neighbouring clouds, 6 minutes before drop.

#### PLATE 2

First experiment. Looking SW. at 12,000 ft. Close-up of cloud top 3 minutes before drop.

#### PLATE 3

First experiment. Looking WSW. at 10,500 ft. Close-up of the eastern end of the infected cloud 20 minutes after the dry ice drop. Precipitation is falling from most of the cloud but the infection has not yet reached the near end.

#### PLATE 4

First experiment. Looking S. at 9400 ft. Taken from about 7 miles north of the cloud. Twenty-three minutes after the dry ice drop the cloud has completely turned to ice and is precipitating heavily.

#### PLATE 5

First experiment. Looking NE. at 9500 ft. Twenty-seven minutes after the dry ice drop the cloud is beginning to dissipate and the precipitation is becoming less heavy.



Cirrocumulus layer

Cloud in  
which dry  
ice was  
dropped



General cumulus  
layer

SMITH.—EXPERIMENTS IN SEEDING CUMULIFORM CLOUD LAYERS WITH DRY ICE

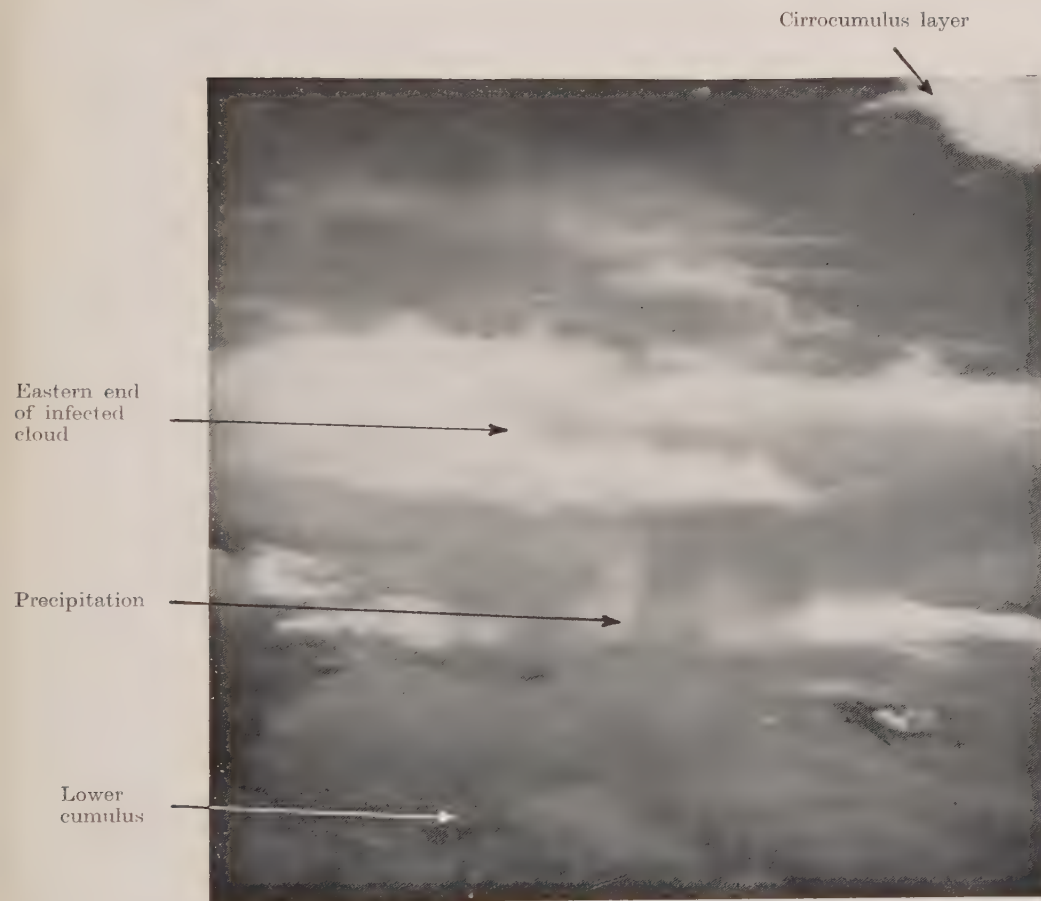




SMITH.—EXPERIMENTS IN SEEDING CUMULIFORM CLOUD LAYERS WITH DRY ICE

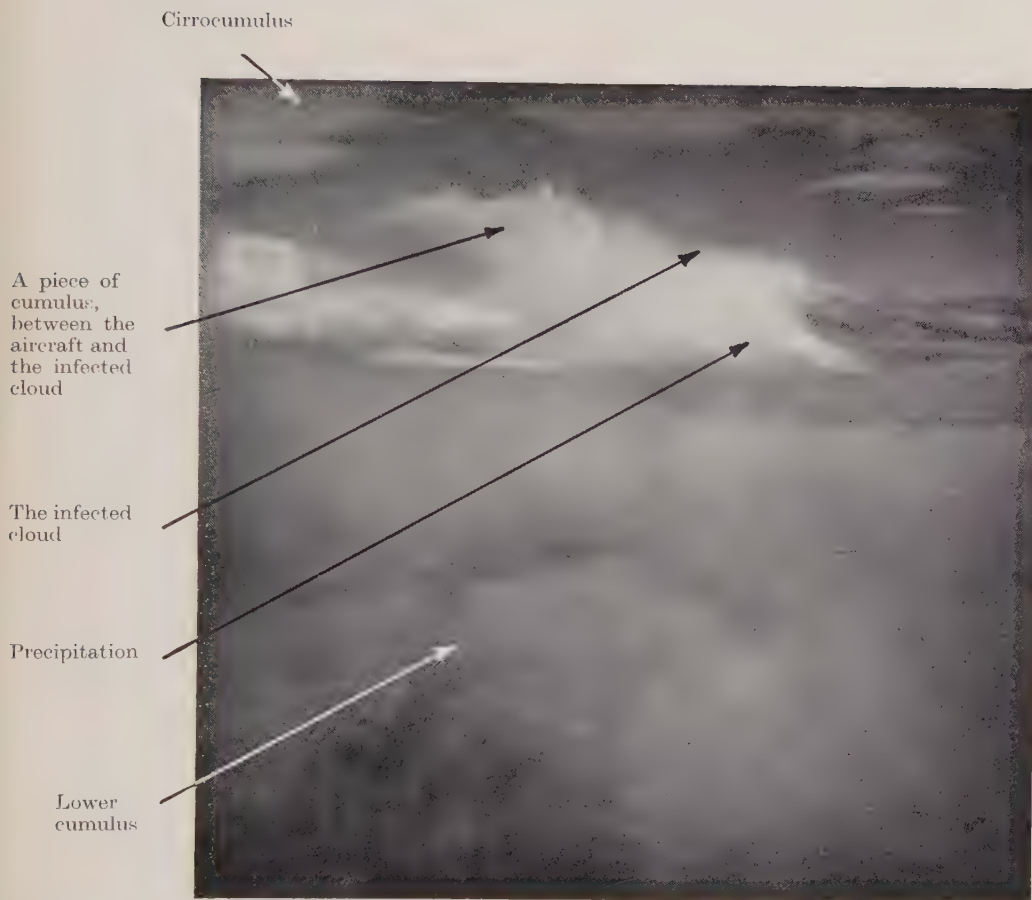






SMITH.—EXPERIMENTS IN SEEDING CUMULIFORM CLOUD LAYERS WITH DRY ICE

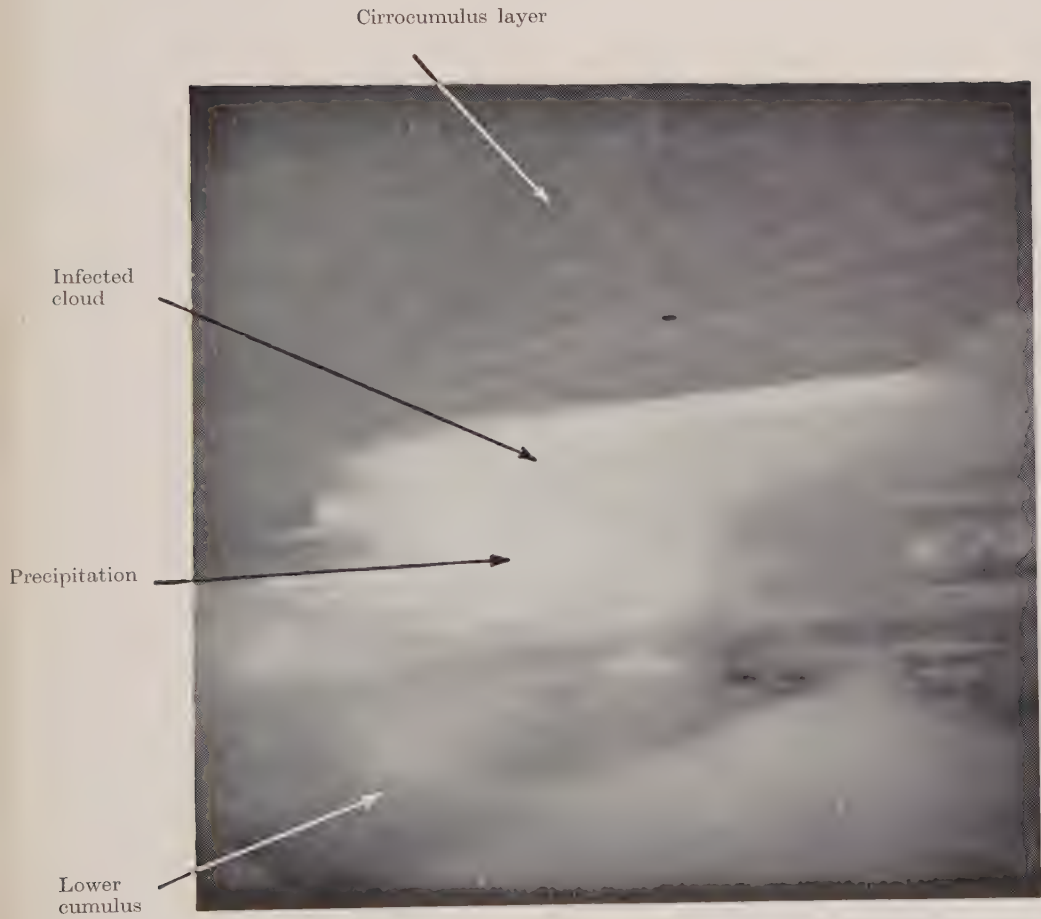




SMITH.—EXPERIMENTS IN SEEDING CUMULIFORM CLOUD LAYERS WITH DRY ICE







SMITH.—EXPERIMENTS IN SEEDING CUMULIFORM CLOUD LAYERS WITH DRY ICE





SMITH.—EXPERIMENTS IN SEEDING CUMULIFORM CLOUD LAYERS WITH DRY ICE

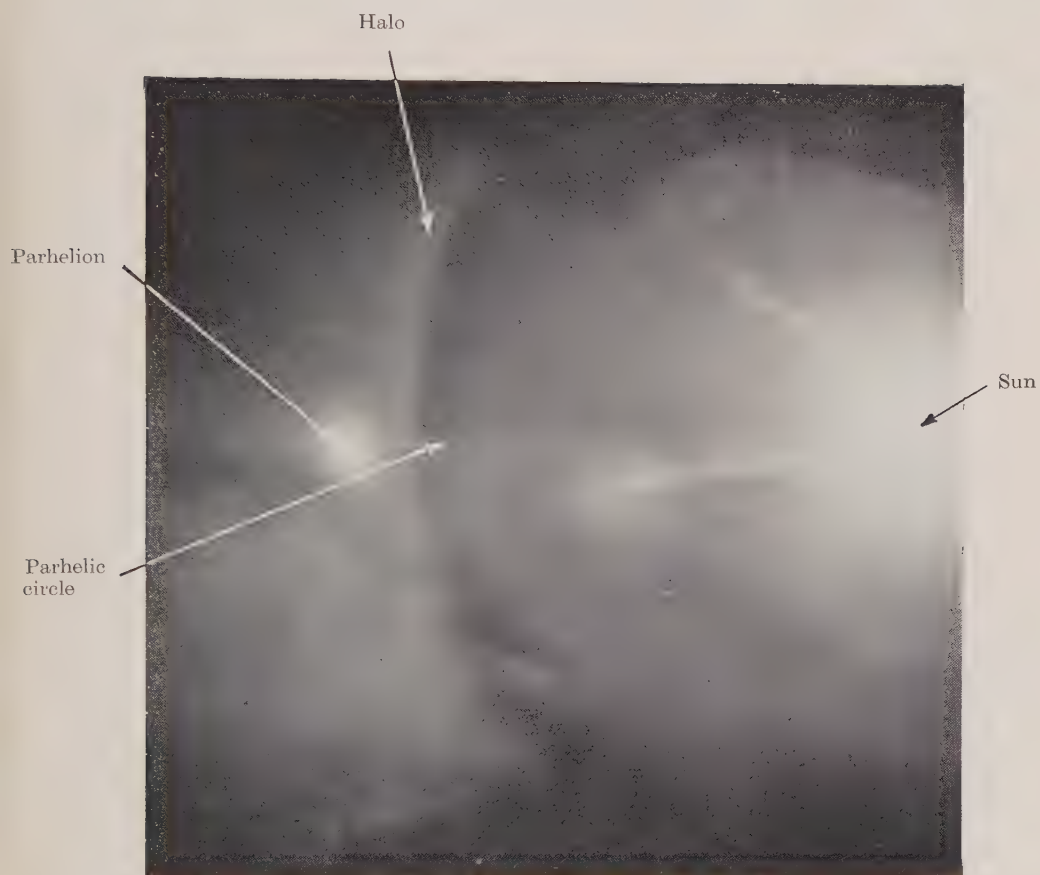






SMITH.—EXPERIMENTS IN SEEDING CUMULIFORM CLOUD LAYERS WITH DRY ICE





SMITH.—EXPERIMENTS IN SEEDING CUMULIFORM CLOUD LAYERS WITH DRY ICE







SMITH.—EXPERIMENTS IN SEEDING CUMULIFORM CLOUD LAYERS WITH DRY ICE





SMITH.—EXPERIMENTS IN SEEDING CUMULIFORM CLOUD LAYERS WITH DRY ICE







SMITH.—EXPERIMENTS IN SEEDING CUMULIFORM CLOUD LAYERS WITH DRY ICE





SMITH.—EXPERIMENTS IN SEEDING CUMULIFORM CLOUD LAYERS WITH DRY ICE

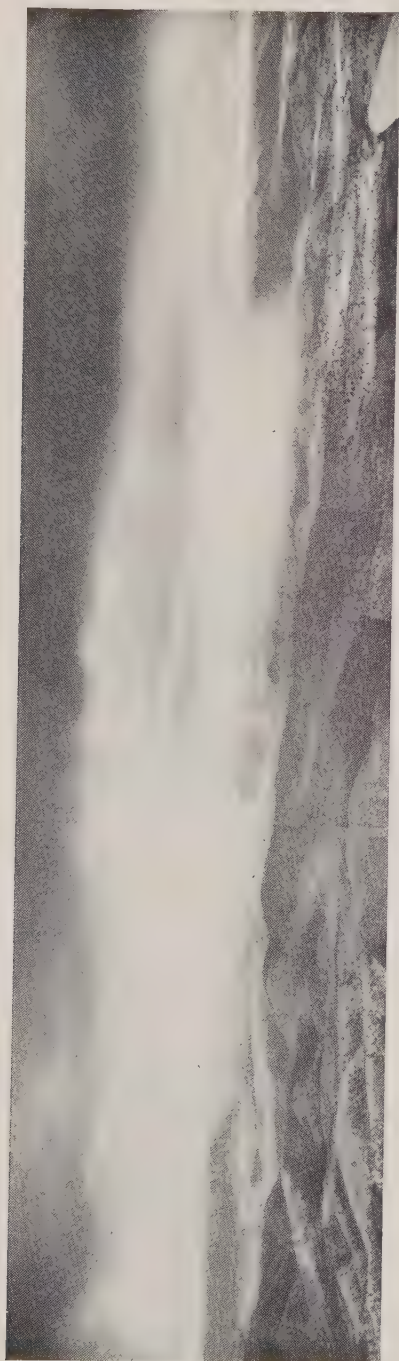






SMITH.—EXPERIMENTS IN SEEDING CUMULIFORM CLOUD LAYERS WITH DRY ICE





SMITH.—EXPERIMENTS IN SEEDING CUMULIFORM CLOUD LAYERS WITH DRY ICE





## PLATE 6

First experiment. Looking S. at 9500 ft. Thirty-seven minutes after the dry ice drop the cloud is disappearing.

## PLATE 7

First experiment. Looking W. at 9500 ft. Photograph taken with a red filter 53 minutes after the dry ice drop. The cloud has almost disappeared. Ice crystals remain and cause the optical phenomena.

## PLATE 8

Second experiment, June 11, 1948. 1407 hours. Looking NW. at 12,000 ft. The sun, 22° halo, parhelson, and parhelic circle. Photograph taken with a red filter.

## PLATE 9

Third experiment, June 17, 1948. 1104 hours. Looking N. at 11,800 ft. The sun, 22° halo, parhelic circle, and parhelson at 25°. Photograph taken with a red filter.

## PLATE 10

Fourth experiment, August 25, 1948. Top of cloud one minute before dry ice drop.

## PLATE 11

Fourth experiment. Twenty-five minutes after dry ice drop. Snow falling to ground. The cloud has become thin and snow is about to stop falling.

## PLATE 12

Fifth experiment, August 25, 1948. Top of cloud one minute before dry ice drop.

## PLATE 13

Fifth experiment. Base of cloud 19 minutes after dry ice drop. Snow beginning to fall, almost reaching ground.

## PLATE 14

Fifth experiment. Base of cloud 38 minutes after dry ice drop. Snow falling over about 30 square miles, melting before reaching ground.

# THE RELATIVE STABILITY OF INTERNAL METAL COMPLEXES

## I. COMPLEXES OF 8-HYDROXYQUINOLINE, SALICYLALDEHYDE, AND ACETYLACETONE

By L. E. MALEY\* and D. P. MELLOR\*

[*Manuscript received November 4, 1948*]

### *Summary*

The stability constants of a series of internal metal complexes derived from 8-hydroxyquinoline, salicylaldehyde, and acetylacetone have been determined by means of an electrometric method.

A quantitative treatment of the chelation of 8-hydroxyquinoline, based on the theory of step equilibria, is given and the modifications required for dealing with the other substances investigated are indicated.

Two important facts emerge from this survey :

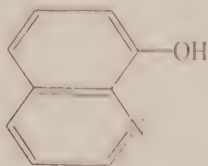
(1) The order of the metals (as regards stability of their complexes) is substantially the same for each of the three chelating substances studied.

(2) The order of the stability of the complexes derived from the three chelating substances is independent of the nature of the coordinating metal, the 8-hydroxyquinoline complex of a metal being more stable than its acetylacetone complex which in turn is more stable than its salicylaldehyde complex.

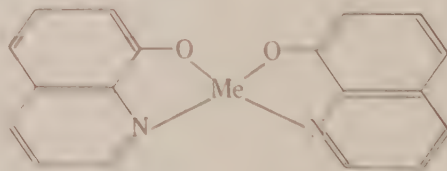
Measurements were made of the stability constants in water-dioxan mixtures and the effect of changing the proportions of the two components in the solvent mixture was studied in several instances. The values of stability constants of internal metal complexes in water-dioxan mixtures are greater than those in water. It is shown that insolubility of internal metal complexes is not, in general, a criterion of their relative stability. The possible significance of stability constants in chemotherapy, biochemistry, and analytical chemistry is briefly discussed.

### I. INTRODUCTION

Many bivalent metal ions readily combine with chelating substances such as 8-hydroxyquinoline (I) to form an internal complex (II)



(I)



(II)

For convenience in later discussion the formula of a substance like (I) will be

\* Department of Chemistry, University of Sydney.

abbreviated to HA and that of the derived metal complex, (II), to  $\text{MeA}_2$ . The stability of (II), defined in relation to the reaction



is measured by the constant  $K_2$  where

$$K_2 = \frac{(\text{MeA}_2)}{(\text{Me})(\text{A}^-)^2}.$$

The reciprocal of  $K_2$ , sometimes called the instability constant, is simply the analogue of the dissociation constants of acids and bases. Investigation of internal metal complexes has, up to the present, been directed mainly to their preparation and structure and to their use in analysis. With one notable exception, very little work has been done in the way of systematic investigation of their physicochemical properties. Calvin and his collaborators(1, 2, 3) have made a detailed study of the stability of internal complexes of copper derived from a wide variety of  $\beta$ -diketones and *o*-hydroxy aromatic aldehydes. Among these compounds a very wide variation in stability was noted.

The present survey was undertaken in order to discover, among other things, (a) whether a similar wide variation in stability occurs with complexes formed by the coordination of a series of metals with any one chelating molecule; and (b) whether complexes formed by different types of chelating molecules with a series of metals showed any systematic variations in stability.

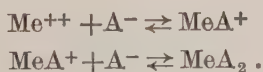
Just as a knowledge of the dissociation constants of acids and bases has done much to systematize our understanding of the behaviour of these substances, so a knowledge of stability constants may be expected to throw light on the behaviour of metal complexes not only in quantitative analysis and in ion exchange separations(4) but also in the complex chemical systems of living organisms(5).

An additional reason for undertaking the present work was to discover whether it might not be possible to correlate stability constants with the bacteriostatic action of substances like 8-hydroxyquinoline believed to exercise this function through their ability to inactivate essential trace elements(6).

Because of the number of equilibria involved, especially those into which the chelating substance itself enters, measurement of the stability constant of an internal metal complex is not quite so direct as the measurement of the corresponding constant of acids and bases. The experimental method used for measuring  $K_2$  is based on the determination of the hydrogen ion concentration of a solution containing known amounts of metal ion, free acid, and the chelating substance. The calculation of the constants from these data, however, is a lengthy process even by what appears to be the most convenient method which is based on the theory of step equilibria.

## II. GENERAL THEORY OF STEP EQUILIBRIA

It has been shown that the formation of a metal complex is an example of a process involving step equilibria(1, 2, 3). In the present instance the steps are



Step equilibria are of wide occurrence and include not only complex formation but acid-base and oxidation-reduction systems. The general theory of step equilibria has been thoroughly worked out by Bjerrum(8) and extensively applied to complex cations (ammines)(8, 9, 10). After briefly outlining the essential features of Bjerrum's theory, its application to a variety of chelating molecules will be given.

The following equilibria correspond to a step system



with formation constants (or step equilibrium constants) defined as follows :

$$\left. \begin{array}{l} k_1 = \frac{(\text{MeA})}{(\text{Me})(\text{A})} \\ k_2 = \frac{(\text{MeA}_2)}{(\text{MeA})(\text{A})} \\ \dots\dots\dots \\ k_N = \frac{(\text{MeA}_N)}{(\text{MeA}_{N-1})(\text{A})} \end{array} \right\} \dots\dots\dots (2)$$

The complexity constant or overall stability constant  $K_N$  is then given by the relation

$$K_N = \frac{(\text{MeA}_N)}{(\text{Me})(\text{A})^N} = k_1 k_2 k_3 \dots k_N \dots\dots\dots (3)$$

In order to simplify the physical analysis of the problem, Bjerrum has introduced several new concepts and relationships. One of these is the quantity  $\bar{n}$  which is defined as the average number of A's attached to Me. Thus

$$\bar{n} = \frac{(\text{MeA}) + 2(\text{MeA}_2) + \dots + N(\text{MeA}_N)}{(\text{Me}) + (\text{MeA}) + (\text{MeA}_2) + \dots + (\text{MeA}_N)} \dots\dots\dots (4)$$

Substituting in equations (2) we obtain the relation

$$\bar{n} = \frac{k_1(\text{A}) + 2k_1k_2\text{A}^2 + \dots}{1 + k_1(\text{A}) + k_1k_2(\text{A})^2 + \dots} \dots\dots\dots (5)$$

This last function is called the formation function of the system and the curve obtained by plotting  $\bar{n}$  against  $\log \frac{1}{(\text{A})}$  is called the formation curve of the system.

Bjerrum has shown that, provided the step constants are not too close together,† there will be equal amounts of  $\text{MeA}_{N-1}$  and  $\text{MeA}_N$  present when

\* Bjerrum(8) describes A as a "ligand", a word meaning a bound atom or group. Since this term is finding increasing use, it will be adopted here. A, Me, and  $\text{MeA}_N$  may carry a positive, negative, or zero charge according to the nature of the system. In the system shortly to be discussed  $\text{Me} = \text{Me}^{++}$ , A is the anion of a bidentate chelate molecule ( $\text{A}^-$ ),  $N=2$ , and  $\text{MeA}_2$  is a complex without a charge (an internal complex).

† It is only in cases where the step constants are very close together that one step reaction interferes with the next. Such cases may be treated by substituting the  $k$  values obtained as above in a further set of equilibrium equations of type (5) (8, p. 35).



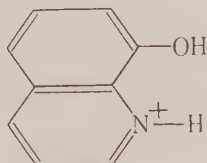
$\bar{n} = N - \frac{1}{2}$  and that the step equilibrium constant  $k_N$  can then be determined by using the formation curve to find the value  $\frac{1}{(A)}$  at  $\bar{n} = N - \frac{1}{2}$ .

The values of  $\log k_1$  and  $\log k_2$  quoted in Tables 1, 2, etc. are the values of  $\log \frac{1}{(A)}$  read from a formation curve at the points  $\bar{n} = \frac{1}{2}$  and  $1\frac{1}{2}$  respectively. The overall stability constant  $K_2 = k_1 k_2$ .

In adapting the general theory to a particular case, the main problem is to derive expressions for  $\bar{n}$  and  $\log \frac{1}{(A)}$  and, as will be seen from the treatment to follow, the manner of doing this depends on the nature of A and the equilibria into which it enters.

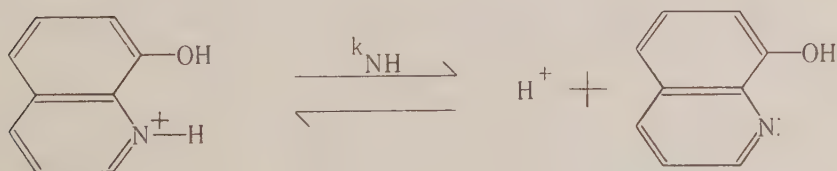
### III. THEORY OF THE METHOD OF CALCULATING THE STABILITY CONSTANTS OF COMPLEXES OF 8-HYDROXYQUINOLINE

To obtain a wide range in the values of  $\bar{n}$  and  $\log \frac{1}{(A)}$  in as convenient a manner as possible, each determination of  $K_2$  was carried out as a potentiometric titration with standard alkali. The solution to be titrated (100 ml.) contained  $T_{Me}$  moles per litre of Me,  $T_L$  moles of 8-hydroxyquinoline, and E moles per litre of free perchloric acid. Standard alkali (such as would produce a concentration of Na moles per litre in the solution) was added and the resulting hydrogen ion concentration ( $H^+$ ) measured by means of a glass electrode. In an acid solution, 8-hydroxyquinoline reacts to form the positively charged ion:



Two dissociation constants of 8-hydroxyquinoline must therefore be taken into account:

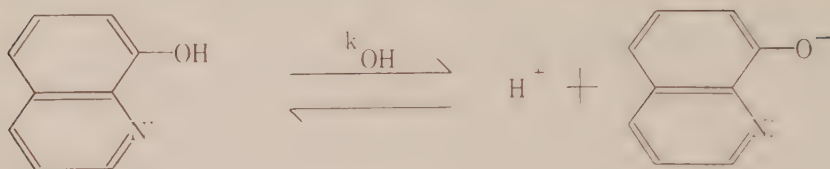
- (1) The dissociation of the N-H group



for which

$$k_{NH} = \frac{(HA)(H^+)}{(HAH^+)}$$

## (2) The dissociation of the O-H group

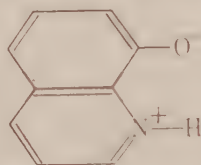


for which

$$k_{OH} = \frac{(A^-)(H^+)}{(HA)}$$

Owing to the fact that the  $pK_a$  of the OH group is much greater than that

of the NH group no zwitterion  $-AH^+$ , i.e.



is formed.

*Derivation of Expressions for  $\bar{n}$  and  $\text{Log } \frac{1}{(A^-)}$*

The derivation will be given for an arbitrary volume of solution and all concentrations will be expressed in moles per litre. The symbols used are defined as follows:

Let  $(T_{Me})$  = total metal combined and uncombined with ligand.

$(T_L)$  = total ligand in all forms.

$(T_H)$  = total hydrogen ionized and potentially ionizable hydrogen.

$$\bar{n} = \frac{\text{ligand attached to metal}}{(T_{Me})}$$

$$n_A = \frac{\text{total potentially ionizable hydrogen atoms bound to ligand}}{\text{total ligand not bound to metal}}$$

$$\alpha = \frac{\text{ligand uncombined with either hydrogen or metal}}{\text{total ligand not bound to metal}}$$

$(E)$  = the initial concentration of added perchloric acid.

$(E - Na)$  = the concentration of excess perchloric acid remaining after the addition of NaOH.

$$(T_{Me}) = (Me^{++}) + (Me^+A) + (MeA_2) \dots\dots\dots (6)$$

$$\bar{n} = \frac{(MeA^+) + 2(MeA_2)}{(T_{Me})} \dots\dots\dots (7)$$

$$(T_L) = (A^-) + (HA) + (HAH^+) + \bar{n} \cdot (T_{Me}) \dots\dots\dots (8)$$

$$(T_H) = (H^+) + (HA) + 2(HAH^+) \dots\dots\dots (9)$$

$$\bar{n}_A = \frac{(HA) + 2(HAH^+)}{(A^-) + (HA) + (HAH^+)} \dots\dots\dots (10)$$

$$\alpha = \frac{(A^-)}{(A^-) + (HA) + (HAH^+)} \dots\dots\dots (11)$$

On eliminating  $(A^-) + (HA) + (HAH^+)$  from  $(T_L)$  and  $\bar{n}_A$ , and then substituting for  $T_H$  we have

$$(T_L) = \frac{(T_H) - (H^+)}{\bar{n}_A} + \bar{n} \cdot (T_{Me}) \dots\dots\dots (12)$$

Since 8-hydroxyquinoline was added initially in the form HA, that is, as a monobasic acid, an equivalent expression for  $(T_H)$  is derived as follows:

$$(T_H) = (H^+) \text{ (from perchloric acid)} + (H^+) \text{ (from water)} + (H^+) \text{ (from ligand)} + (HA) + 2(HAH^+) \text{ and } (H^+) \text{ (from water)} = (OH^-) \text{ (from water),}$$

$$\text{therefore } (T_H) = (E - Na) + \frac{K_w}{(H^+)} + (T_L) \dots\dots\dots (13)$$

$$\text{therefore } \bar{n} = \frac{(T_L) \cdot \left(1 - \frac{1}{\bar{n}_A}\right) - \left\{ \frac{(E - Na) + \frac{K_w}{(H^+)} - (H^+)}{\bar{n}_A} \right\}}{(T_{Me})} \dots\dots\dots (14)$$

$$\text{where } \bar{n}_A = \frac{1 + 2 \frac{(H^+)}{k_{NH}}}{1 + \frac{k_{OH}}{(H^+)} + \frac{(H^+)}{k_{NH}}} \dots\dots\dots (15)$$

An expression for  $\log \frac{1}{(A^-)}$  may be derived in either of two ways. The method used will be described first.

$$(A^-) = \frac{k_{OH}(HA)}{(H^+)}$$

$$\text{therefore } \log \frac{1}{(A^-)} = \log (H^+) - \log k_{OH} - \log (HA).$$

The first two terms on the right-hand side are known from experiment; the third must be calculated.

$$\text{Let } (T_U) = \text{total ligand uncombined with metal} \\ = (A^-) + (HA) + (HAH^+)$$

$$(HA) = (T_U) \{1 - \text{fraction of } (T_U) \text{ in form } (A^-) - \text{fraction of } (T_U) \text{ in form } (HAH^+)\}.$$

The percentage of  $(T_U)$  and therefore the fraction of  $(T_U)$  in the different forms may be calculated by means of Henderson's equation:

$$\text{Percentage of a base in dissociated form} = \frac{100 \times \text{antilog } (pK_a - pH)}{1 + \text{antilog } (pK_a - pH)}$$

$$\text{Percentage of an acid in dissociated form} = \frac{100}{1 + \text{antilog } (pK_a - pH)}$$

A curve showing percentage of an acid in the ionized form as a function of  $(pK_a - pH)$  is shown in Figure 1. The percentage of  $T_U$  as  $(A^-)$  and  $(HAH^+)$  may be read directly from this curve.

$$\text{Log (HA)} = \text{log } (T_L - n \cdot T_{Me}) \{1 - \text{fraction of } (T_U) \text{ in form } (A^-) - \text{fraction of } (T_U) \text{ in form } (HAH^+)\} \dots\dots\dots (16)$$

It is important to note that all  $pK$  values are expressed on the  $pK_a$  scale and that  $(HAH^+)$  is the undissociated form of the acid, the dissociated form being  $HA$ . Care must accordingly be taken with the signs of  $(pK_a - pH)$  to see that they give the correct fractions of the acid and base forms present.

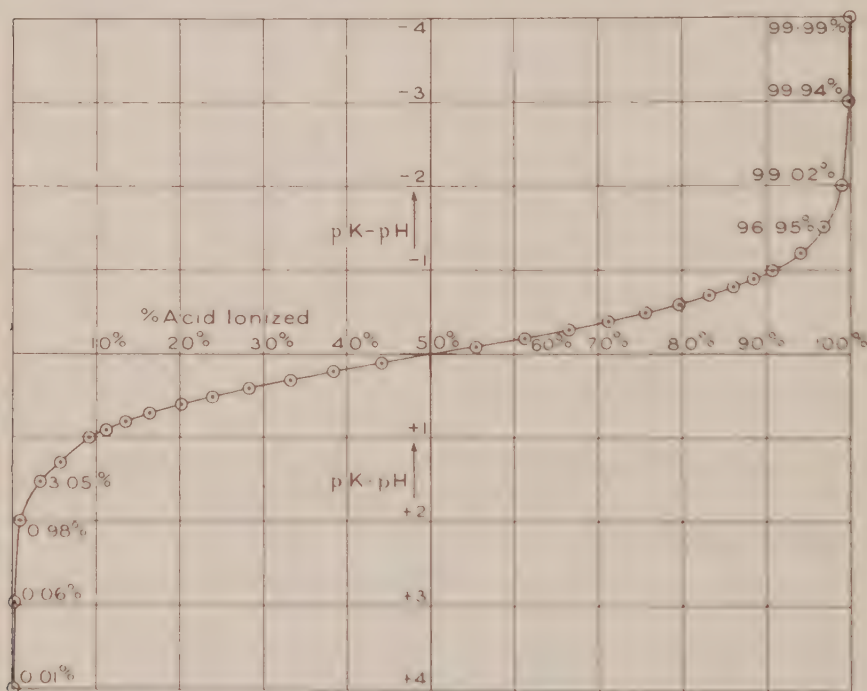


Fig. 1. Graph of Henderson's equation showing the per cent. acid ionized as a function of  $pK_a - pH$ .

Alternatively  $\log \frac{1}{(A^-)}$  may be found from equations of  $\alpha$  and  $n_A$ .

$$(A^-) = \frac{\alpha}{n_A} \{ (HA) + 2(HAH^+) \} \dots\dots\dots (17)$$

where

$$\alpha = \frac{1}{1 + \frac{(H^+)}{k_{OH}} + \frac{(H^+)}{k_{OH}} + \frac{(H^+)}{k_{NH}}}$$

Though at first sight the second method might appear simpler, calculations by this method are in fact longer.



## IV. EXPERIMENTAL DETAILS

(i) *Reagents*.—Owing to the insolubility of 8-hydroxyquinoline-metal complexes in water and because of the need to work with a one-phase system, it was necessary to make all measurements in a 70 per cent. dioxan-water mixture. This mixture was found capable of dissolving sufficient of the least soluble complexes studied. With the use of a 6 ft. fractionating column, dioxan was purified by the Weissberger method(11). Reagents such as perchloric acid, 8-hydroxyquinoline (m.p. 75° C.), and carbonate free sodium hydroxide were of A.R. grade.

Metal perchlorates were not isolated but were made up as stock solutions by dissolving the metal in excess perchloric acid, the amount of the excess acid being controlled so that when the solution was made up to 100 ml. the pH produced was approximately 1.7. This figure was chosen because it was found that little or no chelation occurred in solutions at or below this pH. Concentrations employed in the solutions titrated were: metal perchlorate 0.001 M, 8-hydroxyquinoline 0.002 M. Alkali additions were made from a microburette, and during the titration the solution was thoroughly stirred by means of a mechanical stirrer. The titration vessel was so designed that where necessary, the whole operation could be carried out under purified nitrogen and with the addition of oxygen free caustic soda. All titrations were made in duplicate at  $25 \pm 0.02^\circ \text{C}$ . Excess perchloric acid was determined potentiometrically by a separate titration using the same metal perchlorate solution but with the 8-hydroxyquinoline omitted.

(ii) *Measurement of pH*.—All pH measurements were made with a glass electrode (Leeds and Northrop pH meter: model 7663-A1). The instrument was standardized with two standard buffers of pH 7.00 and 4.00 respectively, supplied by the makers of the instrument. It was found necessary to treat the upper surface of the glass electrode with "dri-film", a treatment which effectively increased the surface insulation to over 100 megohms. The pH readings at given caustic soda additions in duplicate titrations were reproducible with a maximum variation of  $\pm 0.01$  pH unit.

(iii) *Dissociation Constants of 8-Hydroxyquinoline*.—Since pK values of a substance depend on the medium in which they are measured it was necessary to make a determination of the values for 8-hydroxyquinoline in a 70 per cent. dioxan-water mixture. This was done in the usual way by a potentiometric titration of a known weight with standard sodium hydroxide and perchloric acid. The values obtained at 25° C. were:  $\text{pK}_{\text{OH}} = 12.33$ ;  $\text{pK}_{\text{NH}} = 3.18$  both on the acid scale.

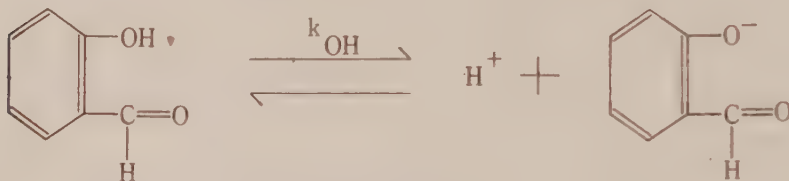
It was necessary in all titrations to make concentration corrections to allow for the increase in volume arising from added standard alkali.

## V. STABILITY CONSTANTS OF SALICYLALDEHYDE AND ACETYLACETONE METAL COMPLEXES

Expressions for determining stability of chelate molecules such as these which have one dissociable hydrogen atom can be derived by the method outlined by Calvin and Wilson(1).

Monoacidic chelating molecules of this type can, however, also be treated quantitatively as follows:

Let the monoacidic dissociation



be denoted by

$$k_{OH} = \frac{(A^-)(H^+)}{(HA)}$$

$$(T_L) = (A^-) + (HA) + \bar{n} \cdot (T_{Me}); \quad (T_H) = (H^+) + (HA)$$

$$(T_{Me}) = (Me^{++}) + (MeA^+) + (MeA_2); \quad \bar{n}_A = \frac{(HA)}{(A^-) + (HA)}$$

also

$$(T_H) = (\text{the concentration of excess acid}) + (T_L)$$

therefore

$$\bar{n} = \frac{(T_L) \left( 1 - \frac{1}{\bar{n}_A} \right) - \left[ \frac{(E - Na) + \frac{K_w}{(H^+)} - (H^+)}{\bar{n}_A} \right]}{(T_{Me})}$$

where

$$\bar{n}_A = \frac{(H^+)}{k_{OH} + (H^+)}$$

when

$$\bar{n}_A = 1 \quad \text{then} \quad \bar{n} = \frac{(Na - E) + (H^+)}{(T_{Me})}$$

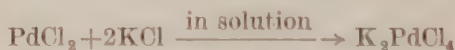
The chelating form here is  $(A^-)$

therefore  $\log \frac{1}{(A^-)} = \log (H^+) - \log k_{OH} - \log (HA)$

where  $\log (HA) = \log (T_L - \bar{n} \cdot T_{Me}) \{1 - \text{fraction of } (T_L) \text{ as } (A^-)\}$ .

Although acetylacetone forms water soluble complexes, both it and salicylaldehyde were studied in 50 per cent. dioxan-water mixtures in order to make a direct comparison of the stability of their complexes. The stability constants were determined in the same way as those for 8-hydroxyquinoline.

A slight variation was introduced in measuring the constants for palladium complexes. Instead of the perchlorate it was necessary to use  $K_2PdCl_4$  which was prepared as follows:



and the final product recrystallized from water. Since this compound, when treated with the monoacidic chelating molecules, liberates two hydrogen ions, it can be treated in the usual way.

The step constants quoted are for the reactions as indicated



The overall reaction is therefore



Solutions containing  $\text{Mn}^{++}$ ,  $\text{Fe}^{++}$ , and  $\text{Co}^{++}$  were titrated in an atmosphere of nitrogen.

*Note on Calculation of Results.*—In calculating the value of the expressions  $\bar{n}$  and  $\log \frac{1}{(\text{A}^-)}$  the instrument readings of pH were used as a measure of hydrogen ion activities but no attempt was made to take into account activity coefficients of the bivalent metal ions and ligand. The bivalent metal ion concentrations used were M/1000 and for most cases, chelation was begun and virtually completed during the addition of 0.2 ml. of normal alkali. The ionic strength of the solution (100 ml.) which, after neutralization of excess acid, was approximately M/50 with respect to  $\text{NaClO}_4$ , did not change appreciably during the addition of the 0.2 ml. of normal alkali. Since, in any series of measurements with one chelating substance, conditions were kept roughly constant except for the variation of the nature of the bivalent metal ion, it is believed that the values of  $K_2$  do indicate relative stabilities.

## VI. RESULTS

Measurements of constants are summarized in Tables 2, 3, and 4, where the metals are arranged in order of decreasing stability of their complexes. A typical set of experimental results is shown in Table 1 and the titration and formation curves of the 8-hydroxyquinoline compounds are shown in Figures 2 and 3. Since the curves for Zn and Co fall very close to one another only that for Co is shown.

*The Effect of a Change of Solvent on Stability.*—In one or two favourable instances it was possible to study the effect on stability of changing the proportions of water and dioxan. It was found that increasing the proportion of dioxan in the solvent not only increases the stability constant of a metal complex but it also increases the  $\text{pK}_a$  value of the chelate itself. Since increasing the proportion of dioxan in the solvent diminishes (12) its dielectric constant, the increase in both  $K_2$  and  $\text{pK}_a$  is not surprising. The magnitude of the changes produced by altering the proportions of dioxan in the solvent is shown in Tables 5 and 6.

## VII. DISCUSSION

### (a) The Order of Metals in the Stability Series\*

One of the most important facts that emerges from this survey is that the order of the metals (as regards stability of their complexes) is substantially

\* Since this work was submitted for publication, two papers have appeared confirming the general order of stability noted here. Using 5-salicylaldehyde-sulphonic acid, Calvin and Melchior (*J. Amer. Chem. Soc.* **70** : 3270 (1948)) have found the order  $\text{Cu} > \text{Ni} > \text{Co} > \text{Zn}$ . A similar order has been noted for quinaldinate by Irving and Williams (*Nature* **162** : 746 (1948)). Both these groups of workers have found independently a correlation between stability and the second ionization potential corresponding to the process  $\text{Me}^+ \rightarrow \text{Me}^{++} + e$  of the metal concerned.

TABLE 1

THE FORMATION CURVE DATA FOR THE CHELATION OF NICKEL WITH  
8-HYDROXYQUINOLINE AT 25° C.

Nickel = 0.001 M, 8-Hydroxyquinoline = 0.002 M

Excess Perchloric Acid = 3.505 ml. of 1.0885 N NaOH

pH	Titration 1.0885 N NaOH (ml.)	$\bar{n}$	$\text{Log } \frac{1}{(A^-)}$
3.57	3.47	0.37	11.71
3.64	3.49	0.47	11.68
3.67	3.50	0.53	11.63
3.75	3.52	0.63	11.56
3.79	3.53	0.70	11.54
4.01	3.56	0.85	11.27
4.31	3.59	1.04	11.08
4.62	3.61	1.20	10.83
5.07	3.63	1.42	10.51
5.30	3.64	1.48	10.33
5.40	3.64	1.53	10.28
6.12	3.67	1.80	9.94
6.80	3.68	1.91	9.57

TABLE 2

STABILITY CONSTANTS OF METALLIC 8-HYDROXYQUINOLINE COMPLEXES IN 70 PER CENT. WATER-  
DIOXAN AT 25° C.

Acid Dissociation Constants of 8-Hydroxyquinoline:  $\text{pK}_{a(\text{OH})} = 12.33$ ;  $\text{pK}_{a(\text{NH})} = 3.18$

Metal	Metal Concentration	Ligand Concentration	$\text{Log } k_1$	$\text{Log } k_2$	$\text{Log } (k_1/k_2)$	$\text{Log } K_2$
Cu .. ..	0.001 M	0.002 M	15.0, .1	14.0, .1	1.0	29.0
Ni .. ..	0.001 M	0.002 M	11.65	10.35	1.3	22.00
Zn .. ..	0.001 M	0.002 M	10.91	9.90	1.01	20.81
Co <sup>II</sup> .. ..	0.001 M	0.002 M	10.85	9.70	1.15	20.55
Fe <sup>II*</sup> .. ..	0.001 M	0.005 M	9.83	9.01	0.82	18.84
Mg .. ..	$\left\{ \begin{array}{l} 0.001 \text{ M} \\ 0.0025 \text{ M} \end{array} \right.$	0.02 M	0.88	5.96	0.92	12.84

\* Ferrous 8-hydroxyquinoline, unless kept under reducing conditions, rapidly oxidizes to the ferric state. It was maintained in the ferrous form by using the reducing agent platinum black in a stream of hydrogen, and under these conditions only two molecules of 8-hydroxyquinoline combined with the ferrous ion.

The platinum black was prepared from  $\text{PtCl}_2$  and alkaline hydrazine sulphate, filtered, and activated with hydrogen before use.



TABLE 3

THE STABILITY CONSTANTS OF METAL COMPLEXES OF SALICYLALDEHYDE IN 50 PER CENT. WATER-DIOXAN AT 25° C.

$pK_a$  of Salicylaldehyde in 50 Per Cent. Water-Dioxan=9.5 at 25° C.

Metal	Metal Concentration	Ligand Concentration	Log $k_1$	Log $k_2$	Log $(k_1/k_2)$	Log $K_2$
Pd .. ..	0.0005 M	0.01 M	7.74	7.03	0.71	14.77
Cu .. ..	$\left\{ \begin{array}{l} 0.001 \text{ M} \\ 0.0025 \text{ M} \end{array} \right.$	0.02 M	7.40	5.91	1.49	13.31
Ni .. ..	0.001 M	0.02 M	5.22	3.97	1.25	9.19
Pb .. ..	0.001 M	0.02 M	5.06	4.04	1.02	9.10
Co <sup>II</sup> .. ..	0.001 M	0.02 M	4.67	3.63	1.04	8.30
Zn .. ..	0.001 M	0.02 M	4.50	3.60	0.90	8.10
Cd .. ..	0.001 M	0.02 M	4.62	3.14	1.48	7.76
Fe <sup>II</sup> .. ..	0.001 M	0.02 M	4.22	3.40	0.82	7.62
Mn <sup>II</sup> .. ..	0.001 M	0.02 M	3.73	3.06	0.67	6.79
Mg .. ..	0.001 M	0.02 M	3.69	3.11	0.58	6.80

the same for each of the three chelating substances studied (see Tables 2, 3, and 4). A similar order has been found also with complexes of 8-hydroxy-quinoline-5-sulphonic acid and a series of amino acids including histidine which will be discussed in Part II of this series.

The only interchanges in this order of stability occur between neighbouring metals such as zinc and cobalt when the stability constants are close together.

The order of the metals, with the chelating molecules studied, in decreasing order of stability, is

Pd<sup>II</sup>, Cu<sup>II</sup>, Ni<sup>II</sup>, Pb<sup>II</sup>, (Co<sup>II</sup>, Zn<sup>II</sup>), Cd<sup>II</sup>, Fe<sup>II</sup>, Mn<sup>II</sup>, Mg<sup>II</sup>.

TABLE 4

STABILITY CONSTANTS OF METALLIC ACETYLACETONE COMPLEXES IN 50 PER CENT. WATER-DIOXAN AT 25° C.

$pK_a$  of Acetylacetone in 50 Per Cent. Water-Dioxan=9.70 at 25° C.

Metal	Metal Concentration	Ligand Concentration	Log $k_1$	Log $k_2$	Log $(k_1/k_2)$	Log $K_2$
Pd .. ..	0.001 M	0.02 M	8.71	8.13	0.58	16.84
Ni .. ..	0.001 M	0.02 M	6.82	5.24	1.58	12.06
Co .. ..	0.001 M	0.02 M	6.30	4.88	1.42	11.18

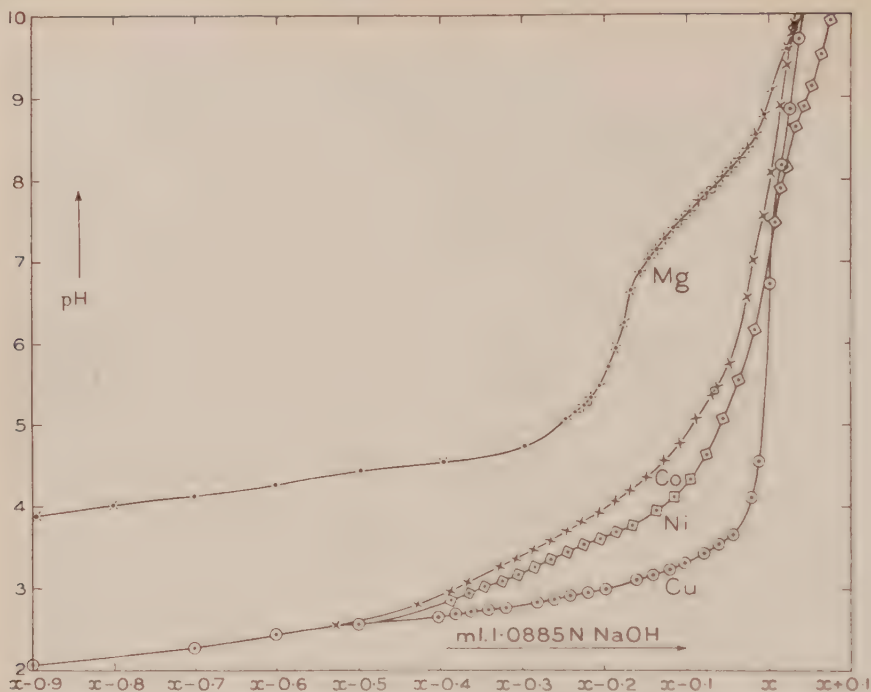


Fig. 2.—Titration curves for the chelation of 8-hydroxyquinoline with different metals.

Each titration was made on a solution with a different amount of excess acid.

The curves have been so plotted that the end points all fall in the same position to give a direct comparison of the stability ranges of the metal complexes.

The titration values of the excess acid were:

3.9205 ml. 1.0885 N NaOH with Copper

3.505       "       "       "       Nickel

3.149       "       "       "       Cobalt

3.215       "       "       "       Magnesium.

The quantity  $x$  is the NaOH required to neutralize the excess acid together with the 0.0002 moles of hydrogen ion liberated as a result of chelation.

E.g. for copper  $x = \left| 3.9205 - \left( \frac{0.2}{1.0885} \right) \right|$  ml. of 1.0885 N NaOH.

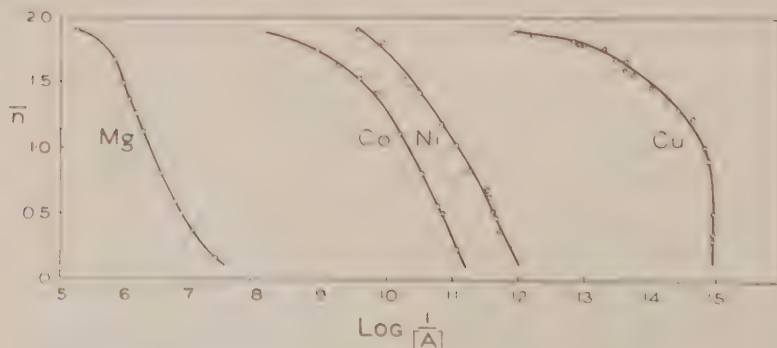


Fig. 3.—The formation curves of 8-hydroxyquinoline metal complexes.

Ethylenediamine complexes have already been shown(13) to fall in the same order. It is of interest to speculate regarding the factors which determine the relative positions of the metals in the series. There is, first of all, a rough parallelism between the stability series and the e.m.f. series.

Stability series	..	..	Pd <sup>II</sup>	Cu <sup>II</sup>	Ni <sup>II</sup>	Pb <sup>II</sup>	Co <sup>II</sup>	Zn <sup>II</sup>	Cd <sup>II</sup>	Fe <sup>II</sup>	Mn <sup>II</sup>	Mg <sup>II</sup>
E.m.f. series	..	..	Pd <sup>II</sup>	Cu <sup>II</sup>	Pb <sup>II</sup>	Ni <sup>II</sup>	Co <sup>II</sup>	Cd <sup>II</sup>	Fe <sup>II</sup>	Zn <sup>II</sup>	Mn <sup>II</sup>	Mg <sup>II</sup>

It will be noticed, however, that the positions of Pb<sup>II</sup> and Zn<sup>II</sup> in the two series are very different.

TABLE 5

THE STABILITY CONSTANTS FOR NICKEL ACETYLACETONE IN DIFFERENT DIOXAN-WATER MIXTURES AT 25° C.

Solvent	Log $k_{OH}$	Metal Concentration	Ligand Concentration	Log $k_1$	Log $k_2$	Log $K_2$
Water ..	8.75	0.001 M	0.02 M	5.92	4.57	10.49
Dioxan 20%..	8.94	0.001 M	0.02 M	6.12	4.79	10.91
Dioxan 50%..	9.70	0.001 M	0.02 M	6.82	5.24	12.06

TABLE 6

THE STABILITY CONSTANTS FOR MAGNESIUM SALICYLALDEHYDE IN DIFFERENT DIOXAN-WATER MIXTURES AT 25° C.

Solvent	Log $k_{OH}$	Metal Concentration	Ligand Concentration	Log $k_1$	Log $k_2$	Log $K_2$
Dioxan 10%..	8.40	0.0005 M	0.01 M	3.44	2.89	6.33
Dioxan 30%..	8.80	0.001 M	0.02 M	3.46	2.93	6.39
Dioxan 50%..	9.50	0.001 M	0.02 M	3.69	3.11	6.80

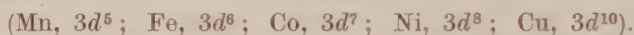
Secondly, electronegativity is a factor that might be expected to play some part in determining stability. If this were so and if electronegativity is accurately defined by the relation deduced by Gordy(20),\* we should expect to find a

\* Electronegativity  $= x = 0.31 \left( \frac{n+1}{r} \right) + 0.5$  where  $n$  = the number of valency electrons (which is 2 in the case here studied) and  $r$  = covalent radius of the atom.

correlation between covalent radii of the metal atom and stability, but no such correlation has been found.

Thirdly, basicity of metals might be expected to play some part in determining stability. For the extreme members of the stability series, there is a rough parallelism between stability and basicity (the more basic the metal, the less stable its complex) but because of difficulty (14) in assigning numerical values to the basicity of metals, no definite statement can be made at present relating stability and basicity.

It is worth noting that metals forming square ( $dsp^2$ ) bonds,  $Cu^{II}$  and  $Pd^{II}$ , head the stability series and that in the transition series of elements investigated, the stability of the complex increases as the electronic  $3d$  orbitals of the corresponding metal atom become more completely filled:



It is interesting to note that a similar effect has been observed(15) with the fifteen rare earth metals in the series La to Lu where the relative stability of

TABLE 7

OVERALL STABILITY CONSTANTS OF A SERIES OF METALS WITH DIFFERENT CHELATING SUBSTANCES

Chelate	Metal				Solvent
	Cu	Ni	Co	Mg	
8-Hydroxyquinoline .. ..	29.0	22.0	20.5	12.8	Dioxan 70%
Acetylacetone .. ..	17.1(1)	12.0	11.2	9.5	Dioxan 50%
Salicylaldehyde .. ..	13.3	9.9	8.3	6.8	Dioxan 50%

their citrate complexes which is the inverse of their relative adsorbability on the exchange columns also increases as the inner  $4f$  shell becomes more completely filled.

### (b) The Effect of the Nature of the Ligand on Stability

A second regularity in the order of stability constants can best be illustrated by referring to Table 7, where it can be seen that the metal complexes of the three chelating substances listed always fall into the same order.

For any given metal, the complex derived from 8-hydroxyquinoline is more stable than that derived from acetylacetone, which is more stable again than the salicylaldehyde complex.

In regard to the last two compounds it is interesting to note that the resonance energy of the chelate ring in acetylacetone is greater than that in salicylaldehyde and that greater resonance energy coincides with greater stability.

Calvin(1) found this relationship to hold among copper complexes and it is evident that, as might be expected, it applies to other metals besides copper.

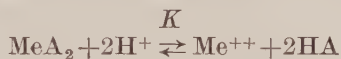


(c) *Solubility and Stability of Internal Metal Complexes*

In recent work on the bacteriostatic action of chelating compounds Albert and his collaborators(6, 7) have used solubility at a specified pH as a measure of the stability of a complex. Though this procedure may be justified in some instances, it is by no means safe to apply it generally. This is evident from some recent measurements made in the laboratory which show that complexes having high stability constants may also be quite soluble in water. Complexes of 8-hydroxyquinoline-5-sulphonic acid and  $\alpha$ -amino acids fall into this class.

Any attempt to find a relation between solubility and stability is complicated by the fact that the solubility measured may not be that of the complex as such. For each metal there is a range of pH (formation region) over which it combines reversibly with the chelate molecule. The range varies a great deal from metal to metal; for those forming the less stable complexes it occurs in the alkaline region, for the most stable complexes the range may lie well within the acid region. It is only in solutions where the pH lies above the formation region that one can speak of the solubility of the metal complex. In solutions of pH below this region, metal complexes break down into metal ions and acidic forms of the ligand, and under these conditions it is the solubility of the chelating molecule and the metal salt which is measured.

Competition between complex and hydrogen ions (considering the case of a ligand which can function as a weak monobasic acid, with a dissociation constant  $K_a$  monoacid) is represented by the equation



The equilibrium constant of this reaction is

$$K = \frac{1}{K_2 \cdot K_a^2}$$

so that with a given chelating acid, the greater the value of  $K_2$  the greater will be the acidity required to decompose the complex.

It is therefore not surprising to find that with some chelating substances the minimum pH's at which their complexes precipitate, are roughly in the same order as their stabilities. An example of this is afforded by the complexes of 8-hydroxyquinoline, the minimum pH's for the precipitation(16) of which are set out in Table 8.

TABLE 8

MINIMUM pH AT WHICH DIFFERENT METAL COMPLEXES OF 8-HYDROXYQUINOLINE PRECIPITATE

Metal	..	..	..	Cu	Ni	Co	Zn	Cd	Mn
pH ..	..	..	..	2.2	2.8	2.8	2.8	4.0	4.3

However, when a comparison is made between complexes derived from different chelating acids the pH range for stability depends on the value of  $1/(K_2 \cdot K_a^2)$  and this relation may be more sensitive to the  $K_a$  value than to the  $K_2$  value.

The value of  $K$  would give a rough indication of the stability range but the solubility of the complex depends on the lattice forces and solvation energies and these may vary considerably even with one chelating molecule, e.g. the water insoluble nickel and soluble copper dimethylglyoximes.

The following general statement can therefore be made: in the pH range over which metal complexes derived from different chelating molecules can exist as such, there need be no correlation between solubility and stability.

#### (d) *Exchange Reactions and Stability*

If one metal complex is more stable than another of the same type, we would expect to find that under suitable conditions one metal will displace another from a less stable complex. This method has already been used by Pfeiffer, Thielert, and Glaser(17) to study qualitatively the relative stability of complexes of bis-salicylaldehyde-ethylenediimine. From a systematic study of reactions carried out in pyridine solutions Pfeiffer and his collaborators place the metals in the following order:



which with one exception is identical with that in Tables 2, 3, and 4. The different position of iron is almost certainly due to the fact that in one case we are dealing with  $\text{Fe}^{\text{II}}$  and in the other  $\text{Fe}^{\text{III}}$ .

Owing to the fact that Pfeiffer's reactions were carried out in pyridine solutions where the displacing metal ions were undoubtedly in the form of pyridine complexes, it would perhaps be more accurate to speak of them as exchange rather than displacement reactions. It is certainly necessary to regard them as such if their implication is to be understood. A typical reaction is



where sal-en = bis-salicylaldehyde-ethylenediimine, and py = pyridine. This point of view introduces a further complication in so far as the reaction would seem, at first sight, to show that the copper pyridine complex is less stable than the nickel pyridine complex, an order which would not be expected from experience with other series of metal complexes. The question then arises: under what circumstances is one justified in drawing conclusions about relative stabilities from experiments of this kind?

An analysis of the problem shows, paradoxically, that we can be certain that the above copper salicylaldehyde-ethylenediimine complex is more stable than the corresponding nickel complex only if the stability constant for the  $\text{Cu}(\text{py})_6^{++}$  complex is greater than that of  $\text{Ni}(\text{py})_6^{++}$ , that is,  $K_6$  for  $\text{Cu} > \text{Ni}$ . Experiment shows that this latter condition is fulfilled(13).

When comparing stabilities of a series of metal complexes derived from one chelating substance and all reactions are carried out in the one solvent it is convenient to regard them as displacement reactions.

Working in pyridine solutions Hammond\* has verified the order of the stabilities of 8-hydroxyquinoline and salicylaldehyde complexes shown in

\* Unpublished experiments in this laboratory by M. Hammond.

Tables 2 and 3. Not only is this qualitative method of determining relative stabilities useful where the electrometric method is inapplicable, for example, with the glyoximes whose  $pK_a$  values are too small to be determined with any certainty, it is also useful in determining the relative stabilities of different chelate compounds of the one metal. Hammond\* has shown for example that nickel salicylaldehyde-ethylenediimine is more stable than nickel bis-salicylaldehyde.

*(e) The Possible Biochemical Significance of Stability Constants*

Although the constants reported in this paper have been obtained with solvents different from those found in biochemical systems the information may not be without significance for these systems. Constants measured in water solutions will be smaller than the corresponding constants for the solvent mixtures used in the present work but it seems highly probable that the metals will fall in the same order as regards their stability in both water-dioxan and water solutions.

Albert and Rubbo(6, 7) have suggested that the origin of the bacteriostatic action of 8-hydroxyquinoline is to be found in its ability to inactivate an essential trace element through complex formation. After some essential trace element has been tied up in the form of an 8-hydroxyquinoline complex it is conceivable that the subsequent addition of the equivalent amount of a metal able to form a still more stable complex may liberate the essential trace element so enabling it to perform its biochemical function once more.

Subsequent work has shown that this may be an oversimplified picture of the process. The following facts have been established (Rubbo, personal communication) :

- (1) 8-Hydroxyquinoline is bacteriostatic in the presence, but not in the absence, of iron and copper.
- (2) The bacteriostatic effect of the iron-8-hydroxyquinoline system can be nullified by the addition of two equivalents of cobalt but that of the copper-8-hydroxyquinoline system is not nullified by the same amount of cobalt.
- (3) The bacteriostatic action of the copper-8-hydroxyquinoline system can be nullified by the addition of ten equivalents of cobalt. The second of these three facts is consistent with the finding that the copper complex is more stable than the corresponding iron complex, with the cobalt occupying an intermediate position, but the third is inexplicable in terms of relative stability constants alone.

It has been shown(18) that an excess of Cu, Ni, Co, Zn, or Mn can cause iron deficiency chlorosis in plants. It is possible that an explanation of this phenomenon will be found in terms of displacement reactions and the relative stability of the complexes involved. For example, the presence of an excess of copper in a plant would almost certainly displace iron from its ferrous complexes. Iron deficiency chlorosis would presumably result from the inability of copper complexes to perform the biochemical functions of the normal iron complexes.



(f) *Coprecipitation with Internal Metal Complexes*

Though a great deal of attention has been paid to coprecipitation among inorganic compounds, little has been done to investigate its extent among internal metal complexes. The only work that appears to have been done along these lines is that of Griffing, de Vries, and Mellon(19) on nickel bis-dimethylglyoxime. Copper is one of the metals coprecipitated with this complex, although copper itself does not, strangely enough, form a water insoluble compound with dimethylglyoxime. Two reasons may be suggested for the coprecipitation of copper under these conditions. Firstly, the copper dimethylglyoxime complex is almost certainly more stable than the nickel complex, and secondly, the shape and size of the two complexes (both metal atoms are square coordinated) are so similar that mixed crystal formation might easily occur. Other factors such as adsorption and pH at which precipitation occurs are also concerned in coprecipitation, but possible interference from soluble complexes of great stability is one that appears to be worth further investigation.

Part II of this series will deal with the chelation stability of a series of  $\alpha$ -amino acids, 8-hydroxyquinoline-5-sulphonic acid, and histidine.

## VIII. ACKNOWLEDGMENTS

One of us (L.M.) is indebted to the Commonwealth Research Committee of the University of Sydney for a grant which enabled the carrying out of this work.

## IX. REFERENCES

- (1) CALVIN, M., and WILSON, K. W.—*J. Amer. Chem. Soc.* **67**: 2003 (1945).
- (2) DUFFIELD, R. B., and CALVIN, M.—*Ibid.* **68**: 557 (1946).
- (3) CALVIN, M., and BAILES, R. H.—*Ibid.* **68**: 949 (1946).
- (4) TOMPKINS, E., KHYM, J., and COHN, W.—*Ibid.* **69**: 2776 (1947).
- (5) GREEN, D. E.—“Mechanism of Biological Oxidation”. p. 106. (Cambridge Univ. Press, 1941.)
- (6) ALBERT, A.—*Med. J. Aust.* **1**: 245 (1944).
- (7) ALBERT, A., RUBBO, S. D., GOLDAURE, R. J., and BALFOUR, B. G.—*Brit. J. Exp. Path.* **28**: 69 (1947). (Other publications in press.)
- (8) BJERRUM, J.—“Metal Ammine Formation in Aqueous Solutions.” (P. Haase and Son: Copenhagen, 1941.)
- (9) BJERRUM, J., and ANDERSON, R. *Kgl. Danske Videnskab. Selskab. Mat. fys. Med. 22*, 7: 1 (1945).
- (10) CARLSON, G. A., McREYNOLDS, J. P., and VERHOEK, F. H.—*J. Amer. Chem. Soc.* **67**: 1334 (1945).
- (11) WEISSBERGER, A., and PROSKAUER, E.—“Organic Solvents”. p. 140. (Oxford Univ. Press, 1935.)
- (12) HARNED, O.—“The Physical Chemistry of Electrolytic Solutions”. p. 544. (Reinhold Publishing Corp.: New York, 1943.)
- (13) MELLOR, D. P., and MALEY, L.—*Nature* **161**: 436 (1948).
- (14) BRITTON, H.—“Hydrogen Ions”. pp. 2, 35. (Chapman & Hall: London, 1942.)
- (15) KETELLE, B. H., and BOYD, G. E.—*J. Amer. Chem. Soc.* **69**: 2810 (1947).
- (16) MELLAN, L.—“Organic Reagents in Inorganic Analysis”. p. 120. (The Blakiston Co., Philadelphia, 1941.)
- (17) PFEIFFER, P., THIELERT, H., and GLASER, H.—*J. Prakt. Chem.* **152**: 145 (1939).
- (18) MILLIKAN, C. R.—*Nature* **161**: 528 (1948).
- (19) GRIFFING, M., DE VRIES, T., and MELLON, M. G.—*Anal. Chem.* **19**: 654 (1947).
- (20) GORDY, W.—*Phys. Rev.* **69**: 604 (1946).



# THE IODINATION OF AROMATIC COMPOUNDS

## I. THE ACTION OF IODINE ON ANILINE IN THE PRESENCE OF MERCURIC OXIDE

By L. JURD\*

[*Manuscript received October 25, 1948*]

### *Summary*

A satisfactory process for the preparation of *p*-iodo- and 2,4-diiodo-aniline has been developed, aniline being treated with iodine in the presence of mercuric oxide. Acetanilide was also iodinated by this procedure. When lead, calcium, and ferric oxides were employed the yields of *p*-iodo-aniline proved to be only slightly greater than that obtained by the action of iodine alone, whilst zinc oxide lowered the yield considerably.

The influence of added iodides, acids, and mercuric iodide on the rate of reaction of iodine and aniline was investigated.

A theory has been proposed to explain the role of mercuric oxide in effecting the iodination of aniline.

## I. INTRODUCTION

Owing to the reactivity of the amine molecule most of the reagents which have been employed to effect direct iodination of aniline gave poor yields of the iodo-amines (Table 1). Better yields of *p*-iodo-aniline were obtained in an indirect process involving the preparation and hydrolysis of *p*-iodo-acetanilide.

## II. THE IODINATION OF ANILINE IN THE PRESENCE OF MERCURIC OXIDE

It has now been found that equimolecular quantities of iodine and aniline in the presence of 0.75 molecular equivalents of mercuric oxide react in ethanol solution at 5° C. to give 80 per cent. yields of *p*-iodo-aniline. When two or three molecular equivalents of iodine are employed, 60 per cent. yields of 2,4-diiodo-aniline, but no 2,4,6-triiodo-aniline, are obtained. The yields are slightly greater in 95 per cent. ethanol than in absolute alcohol, and both yellow and red mercuric oxide are equally effective.

Lead, calcium, and ferric oxides, in contrast to mercuric oxide, only increase the yield slightly above that obtained in the absence of added oxide, whilst zinc oxide actually decreases the yield (Table 2). However, calcium and zinc oxides were shown to be as effective as mercuric oxide in preventing aniline hydriodide formation (with consequent deactivation of the amine) during the reaction.

Increase in the hydrogen ion and iodide ion concentrations in the reaction mixture decreases the rate and extent of the reaction between iodine and aniline,

\* Massey Agricultural College, Palmerston North, New Zealand.

while the non-electrolyte mercuric iodide has a powerful catalytic effect (Fig. 2). These observations can be readily explained in terms of the theory that the reaction is between a positively polarized iodine atom and the nucleophilic centres

TABLE I  
IODINATION OF ANILINE AND ACETANILIDE

Iodinating Agent	Temperature	Product	Yield (%)	Melting Point	Reference
<i>(a) Aniline</i>					
Iodine monochloride					
1 mole .. ..	Room temp.	4-Iodo-aniline	Very small	Not constant	(1)
2 mole .. ..	5° C.	2,4-Diiodo-aniline	Small	55.5° C.	(2)
3 mole .. ..	Room temp.	2,4,6-Triiodo-aniline	85	165-178°C.	(3)
Chloramine and hydrogen iodide	Room temp.	4-Iodo-aniline			(4)
	100° C.	2,4,6-Triiodo-aniline			
Iodine and lead oxide	100° C.	Indefinite	Very small	Not constant	(5)
Iodine and sodium persulphate ..	Warm	4-Iodo-aniline	4	56° C.	(6)
		2,4-Diiodo-aniline	6	95° C.	
Iodine and sodium bicarbonate ..	12-15° C.	4-Iodo-aniline	75-84	62-63° C.	(7)
Iodine .. ..	20-40° C.	4-Iodo-aniline	45	63° C.	(8)
	80-100° C.	2,4-Diiodo-aniline	30	95° C.	
Iodine and dilute acetic acid ..	15° C.	4-Iodo-aniline	30-40	63° C.	(9)
	70-80° C.	2,4-Diiodo-aniline	Small	95° C.	
<i>(b) Acetanilide</i>					
Iodine monochloride	Room temp.	4-Iodo-acetanilide	90	184° C.	(10)
Chloramine and hydrogen iodide	100° C.	4-Iodo-acetanilide	80		(4)
Iodine and sodium persulphate ..	Warm	4-Iodo-acetanilide	Very small		(6)
Iodine and dilute acetic acid ..	15° C.	Acetanilide	100		(9)

at the *ortho*- and *para*- positions to the amino group. An iodide ion will deactivate the iodine molecule by converting it to the negatively charged  $I^-$  (Robertson and co-workers, 11, 12) while acids will deactivate the aniline molecule by converting it to an anilinium ion (Remick, 13). The catalytic effect of mercuric iodide is probably due to the formation of complexes of the type  $I^+HgI^-$

(or with an  $\text{HgI}_4^-$  negative ion) which by strongly polarizing the iodine molecule increase its activity in electrophilic substitution.

This explanation is similar to that given by Pfeiffer and Wizinger(14) for the catalytic activity of aluminium chloride and ferric bromide in aromatic halogenations.

Acetanilide was also iodinated by means of iodine and mercuric oxide, 40 per cent. yields of *p*-iodo-acetanilide being obtained. Attempts to introduce iodine into the *ortho*-positions of *p*-iodo- and *p*-bromo-acetanilides were abortive, the unchanged starting materials being recovered.

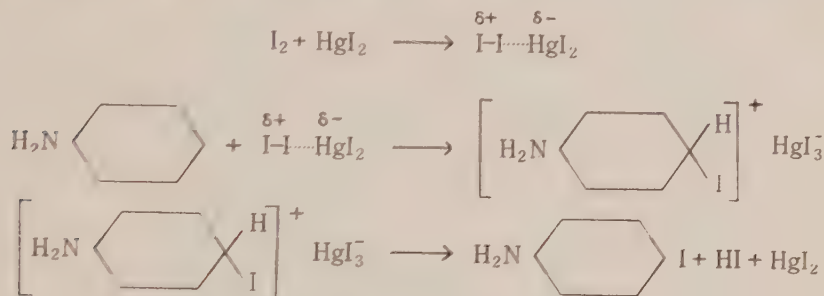


Fig. 1

### III. EXPERIMENTAL

#### (a) Preparation of *p*-Iodo-aniline

(i) A suspension of mercuric oxide (11.3 g.) in an alcoholic solution (20 cc.) of freshly-distilled aniline (6.5 g.) was cooled to 5° C. and iodine (17.6 g.) was then added in several small portions during ten minutes with vigorous agitation between each addition. After shaking for a further ten minutes the reaction mixture was placed in the refrigerator for thirty minutes, the insoluble mercury compounds then being removed by filtration, washed with boiling alcohol (5 cc.) and placed aside for the recovery of mercury. The dark, oily precipitate which separated when the combined filtrate and washings were poured into water (200 cc.) containing potassium iodide (1.5 g.) was extracted with ether. After washing the ethereal solution with aqueous potassium iodide and sodium thiosulphate solutions to remove any remaining mercuric iodide and iodine respectively, the solvent was distilled and the residual oil suspended in water, the mixture made strongly alkaline with sodium hydroxide, and steam-distilled. *p*-Iodo-aniline, in the form of almost white needles, melting at 60° C., separated from the distillate (12.2 g., i.e. 80 per cent. yield). One recrystallization of this product from low-boiling petroleum ether gave long, colourless needles, melting at 62–63° C. (10.9 g.; 72 per cent.).

(ii) The procedures adopted in the preparation of *p*-iodo-aniline by the action of iodine on aniline, alone and in the presence of the oxides of lead, calcium, iron, and zinc were similar to, but less elaborate than, that employed with mercuric oxide since these metals are non-volatile. In the general procedure, iodine (1 mole), the oxide (0.75 mole), and aniline (1 mole) were allowed to react in an alcoholic medium for 1.5 hours. After removing the insoluble metallic compounds, the dark, alcoholic solution was poured into water (250 cc.) containing sodium bisulphite to destroy unreacted iodine. The heavy precipitate which separated was collected, suspended in 5 per cent. sodium hydroxide solution and steam-distilled. *p*-Iodo-aniline separated from the distillate as white needles, melting at 60–61° C. The yields obtained with each of the oxides are given in Table 2.

#### (b) Preparation of 2,4-Diiodo-aniline

Iodine (4.6 g.) was added in four portions during fifteen minutes, with vigorous agitation between each addition, to a suspension of mercuric oxide (2.9 g.) in a solution of *p*-iodo-aniline (4.0 g.) in alcohol (20 cc.), the temperature of the reaction mixture being maintained between

5–10° C. After shaking for a further fifteen minutes and standing for four hours in the refrigerator, the insoluble residue was filtered off, washed with two 10 cc. portions of cold alcohol and allowed to dry (A). The crude diiodo-aniline, which separated as a flocculent precipitate when water (250 cc.) containing a little sodium bisulphite and potassium iodide was added to the alcoholic filtrate, was collected and dried (2.58 g.). On extracting this dark product several times with boiling petroleum ether and allowing the extracts to cool, almost colourless, short needles melting at 95° C., were obtained (2.1 g.). The alcohol-insoluble residue (A), similarly extracted with petroleum ether, yielded brown needles, melting at 94° C. (1.7 g.). These crystalline products were combined and recrystallized from petroleum ether, short needles, only faintly coloured and melting at 95.5° C., being obtained (3.4 g.; 54 per cent.).

TABLE 2

Aniline (Moles)	Iodine (Moles)	Oxide (0.75 Mole)	4-Iodo-aniline Yield. (%)
1.0	1.0		40–1
1.0	1.0	Mercuric oxide	80–1
1.0	1.0	Lead monoxide	46–8
1.0	1.0	Calcium oxide	44
3.0	3.0	Ferric oxide	44–5
1.0	1.0	Zinc oxide	28–31

2,4-Diiodo-aniline was also prepared in 56–60 per cent. yields by the action of iodine (2 moles) and mercuric oxide (1.5 moles) on aniline (1 mole). When three moles of iodine per mole of aniline were employed, no triiodo-aniline was detected among the products of the reaction, the only substance isolated being 2,4-diiodo-aniline in 65 per cent. yield.

#### (c) Preparation of *p*-Iodo-acetanilide

Iodine (9.7 g.) was added gradually during one hour to a vigorously agitated mixture of mercuric oxide (6.0 g.) and acetanilide (5.0 g.) in alcohol (20 cc.) maintained at 5° C. After standing for twenty hours in the refrigerator, the mercury compounds were removed by filtration and washed with ether, the ether washings being added to the alcoholic filtrate. Excess ether was added and the solution extracted with dilute potassium iodide and sodium hydroxide solutions and finally washed with water. The oil which remained on distilling the ether solidified when treated with dilute potassium iodide solution to a white mass, melting at 178° C. (4.0 g.). On crystallizing this crude *p*-iodo-acetanilide from benzene, colourless rhombs, melting at 183° C., were obtained (3.4 g.; 35 per cent.).

#### (d) Rate of Reaction of Iodine and Aniline

To determine the influence of iodides, acids, and mercuric iodide on the rate of reaction, 95 per cent. ethanol solutions of sulphuric acid, hydrogen iodide, and potassium iodide, each containing 0.107 mole per litre, mercuric iodide, 0.027 mole per litre, iodine, 0.215 mole per litre, and of aniline, 0.977 mole per litre, were prepared. Iodine solution (50.0 cc.) and the appropriate volume of the iodide or acid solutions were mixed and the volume made up to 150 cc. with alcohol. After cooling the mixture to 5° C., aniline solution (11.0 cc.) was added from a rapid-flowing pipette. 5 cc. samples of the reaction mixture were withdrawn at intervals, run into an aqueous-acid solution of potassium iodide to stop the reaction and the unrestricted iodine titrated with standard sodium thiosulphate. At the same time, a blank experiment was carried out, an equal



volume of alcohol being added in the place of the aniline solution. The differences in the titrations obtained in each experiment thus gave a measure of the amount of iodine which had reacted. The results obtained are indicated in Figure 2.

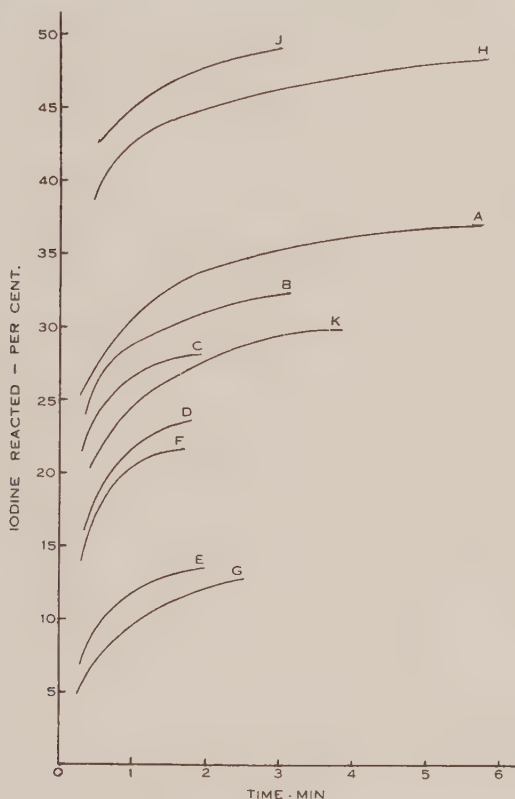


Fig. 2.—Curves illustrating the influence of iodide ions, hydrogen ions, and mercuric iodide on the rate of reaction between equimolecular quantities of iodine and aniline. All additives are expressed in molecular equivalents.

*A*, iodine and aniline alone; *B*, plus sulphuric acid (0.1); *C*, plus sulphuric acid (0.2); *D*, plus potassium iodide (0.2); *E*, plus potassium iodide (0.4); *F*, plus hydrogen iodide (0.2); *G*, plus hydrogen iodide (0.4); *H*, plus mercuric iodide (0.25); *J*, reaction mixture saturated with mercuric iodide; and *K*, reaction mixture saturated with lead iodide.

#### IV. ACKNOWLEDGMENTS

The author wishes to acknowledge his indebtedness to Mr. G. K. Hughes, Department of Organic Chemistry, University of Sydney, for his generous advice and interest in this project, and to Dr. C. R. Barnicoat, Mr. W. McGillivray, and Mr. R. W. Cawley, Department of Biochemistry, Massey Agricultural College, New Zealand, for their interest and cooperation.

## V. REFERENCES

- (1) STENHOUSE, J.—*J. Chem. Soc.* **17**: 327 (1864).
- (2) MICHAEL, A., and NORTON, L. M.—*Ber. dtsh. chem. Ges.* **11**: 107 (1878).
- (3) JACKSON, C. L., and WHITMORE, F. C.—*J. Amer. Chem. Soc.* **37**: 1522 (1915).
- (4) BRADFELD, A., ORTON, K. P., and ROBERTS, I.—*J. Chem. Soc.* **1928**: 782 (1928).
- (5) MILLS, E.—*Ber. dtsh. chem. Ges.* **175-176**: 352 (1875).
- (6) ELBS, K., and VOLK, H.—*J. Prakt. Chem.* **99**: 269 (1919).
- (7) BREWSTER, R. Q. "Organic Syntheses." Coll. Vol. 2, p. 347. (John Wiley and Sons: New York, 1944.)
- (8) HODGSON, H. H., and MARSDEN, E.—*J. Chem. Soc.* **1937**: 1365 (1937).
- (9) MILITZER, W., SMITH, E., and EVANS, E.—*J. Amer. Chem. Soc.* **63**: 436 (1941).
- (10) CHATTAWAY, F. D., and CONSTABLE, A.—*J. Chem. Soc.* **105**: 124 (1914).
- (11) ROBERTSON, P. W., DE LA MARE, P. B. D., and JOHNSTON, W. T. G.—*Ibid.* **1943**: 276 (1943).
- (12) LAMBOURNE, L. J., and ROBERTSON, P. W.—*Ibid.* **1947**: 1167 (1947).
- (13) REMICK, A. E. "Electronic Interpretations of Organic Chemistry". p. 93. (John Wiley and Sons: New York, 1943.)
- (14) PFEIFFER, P., and WIZINGER, R.—*Liebigs Ann.* **461**: 132 (1928).

# STUDIES ON THE LIGNIN OF *EUCALYPTUS REGNANS*

## III. THE ISOLATION OF AN ALKALI LIGNIN FROM SULPHATE PULPING

By J. W. T. MEREWETHER\*

[Manuscript received December 1, 1948]

### Summary

An alkali lignin containing no sulphur has been obtained as a by-product from the sulphate pulping of *Eucalyptus regnans*. Like other alkali lignins it contains hydroxyl groups, both acidic and alcoholic, as well as methoxyl groups.

Acetic anhydride in pyridine yields an octoacetyl derivative which is readily hydrolysed by boiling water to a heptacetyl derivative. In pyridine, benzoyl chloride yields an octobenzoyl derivative while in aqueous alkali it yields a hexabenzoyl compound.

Dimethyl sulphate yields a heptamethyl alkali lignin, diazomethane a hexamethyl derivative, while cold alkaline hydrolysis of the latter gives a pentamethyl derivative, and hot alkaline hydrolysis yields an anhydrotrimethyl alkali lignin.

One of the methoxyl groups formed by methylation is unstable to acetylation by acetic anhydride in pyridine, heptamethyl alkali lignin yielding a hexamethyldiacetyl derivative, hexamethyl alkali lignin a pentamethyltriacetyl derivative, pentamethyl alkali lignin a tetramethyltetracetyl derivative, and anhydrotrimethyl alkali lignin the corresponding anhydrodimethyltetracetyl alkali lignin.

Triphenylchloromethane in pyridine yields a monotrityl derivative.

*p*-Nitrophenylhydrazine gives a di-*p*-nitrophenylhydrazone and phenylhydrazine a phenylhydrazone-osazone.

Potentiometric titration shows two points of inflection and an equivalent weight of 863.

The data are consistent with the empirical formula  $C_{92}H_{104}O_{34}(1754)$  or  $C_{73}H_{84}O_{27}(OCH_3)_{14}(OH)_8C(OH)CO\cdot CH_2OH, CO, COOH$ .

### I. INTRODUCTION

Alkali lignin, that is, lignin isolated with alkali and precipitated by means of mineral acids, has been used as a starting material by many investigators(1), and has been prepared on a commercial scale as a by-product in the manufacture of wood pulp by the soda process (e.g. 2, 3, 4, 5). There is, however, no record of its isolation from the spent liquors of the sulphate pulping process where its formation could be expected from the reaction of the lignin with the sodium hydroxide present. From spruce(6) and *Eucalyptus regnans*(7, 8) the main reaction product has been shown to be a thioglignin, an alkali soluble lignin derivative containing sulphur in the form of a mercaptan group, and presumably formed by the reaction of lignin with sodium hydrosulphide. The present paper deals with the isolation from the sulphate pulping of *Eucalyptus regnans* of an alkali soluble lignin derivative free from sulphur, presumably an alkali lignin.

\* Research Laboratories, Australian Paper Manufacturers Limited, Melbourne.

In a detailed investigation, Marshall, Brauns, and Hibbert(9) found that when wood meal is treated with alkali at 172° C. and the spent liquors acidified with hydrochloric acid the resulting alkali lignin could be separated into two fractions. These two fractions differed considerably in their behaviour to methylating and acetylating reagents, and could be separated by virtue of their different solubilities in 1:10 dioxan-ether.

By applying the technique of Marshall, Brauns, and Hibbert, with suitable modifications, the dioxan-ether mother liquors from the purification of *Eucalyptus regnans* thioglignin(7) yielded a second fraction amounting to 20 per cent. of the crude product. This fraction differed from the thioglignin in that it contained no sulphur and is designated an alkali lignin. Samples prepared from different batches and purified in different ways (precipitating from (a) dioxan into ether-petroleum ether, (b) butanol into petroleum ether, and (c) chloroform into petroleum ether) were all identical as far as could be determined by chemical analysis.

This alkali lignin is a fine powder with a lower softening point (112° C.) than thioglignin (230° C.). It is much lighter in colour, yellow when freshly prepared, and fading, on standing, to a light buff. It is also higher in methoxyl content, 24.7 per cent. compared with 21.1 per cent. methoxyl for thioglignin. Potentiometric titration gives a curve of the same type as that obtained with thioglignin(7), having two points of inflexion—occurring at pH 7.0 and 6.5—and showing an equivalent weight of 863 compared with the value of 665 found for thioglignin.

On treating alkali lignin with acetic anhydride in pyridine there is formed an acetylated alkali lignin with 16.8 per cent. acetyl and 21.1 per cent. methoxyl. However, portion of this acetyl is very loosely bound, and is hydrolysed by boiling water, yielding a compound with 14.8 per cent. acetyl and 21.3 per cent. methoxyl. From these figures a minimum value for the molecular weight can be derived. If the molecular weight of alkali lignin is  $M$  and the molecule contains  $n$  methoxyl groups,  $a$  hydroxyl groups yielding acetates stable to boiling water, and  $b$  hydroxyl groups yielding acetates hydrolysed by boiling water, then

$$\frac{n \times 31.03}{M} = 0.247 \quad \dots \dots \dots (1)$$

$$\frac{n \times 31.03}{M + 42a} = 0.213 \quad \dots \dots \dots (2)$$

$$\frac{n \times 31.03}{M + 42(a + b)} = 0.211 \quad \dots \dots \dots (3)$$

$$\frac{a \times 43.04}{M + 42a} = 0.148 \quad \dots \dots \dots (4)$$

$$\frac{(a + b)43.04}{M + 42(a + b)} = 0.168 \quad \dots \dots \dots (5)$$

From (2) and (4)

$$\begin{aligned} \frac{n \times 31.03}{a \times 43.04} &= \frac{0.213}{0.148} \\ \therefore n/a &= 2/1 \quad \dots \dots \dots (6) \end{aligned}$$



From (3) and (5)

$$\frac{n \times 31.03}{(a+b)43.04} = 0.211$$

$$\therefore n/(a+b) = 7/4 \dots\dots\dots (7)$$

Hence from (6) and (7)

$$n : a : a+b = 14 : 7 : 8$$

$$\text{or } n : a : b = 14 : 7 : 1 \dots\dots\dots (8)$$

Treatment of alkali lignin with diazomethane yields a methylated product with 33.7 per cent. methoxyl. Again, portion of this methoxyl is not strongly bound, and treatment of the methylated product with cold alcoholic potassium hydroxide yields a product with 32.5 per cent. methoxyl. A more completely methylated product is obtained by the simultaneous deacetylation and methylation of the completely acetylated alkali lignin with dimethyl sulphate and 30 per cent. sodium hydroxide, yielding a product with 35.2 per cent. methoxyl. If there are  $c$  hydroxyl groups which react with diazomethane to yield methoxyls stable to cold alcoholic alkali,  $d$  hydroxyls that yield methoxyls hydrolysed by that reagent, and  $e$  hydroxyls that react with dimethyl sulphate but not with diazomethane then

$$\frac{(n+c)31.03}{M+14c} = 0.325 \dots\dots\dots (9)$$

$$\frac{(n+c+d)31.03}{M+14(c+d)} = 0.337 \dots\dots\dots (10)$$

$$\frac{(n+c+d+e)31.03}{M+14(c+d+e)} = 0.352 \dots\dots\dots (11)$$

Substituting in each case for  $M$  from equation (1) we have

$$n/c = 14/5.16 \dots\dots\dots (12)$$

$$n/(c+d) = 7/3 \dots\dots\dots (13)$$

$$\frac{n}{c+d+e} = \frac{2}{1} \dots\dots\dots (14)$$

Hence

$$n : c : d : e = 14 : 5 : 1 : 1 \dots\dots\dots (15)$$

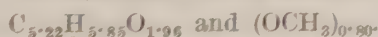
From equations (8) and (15), it is seen that the molecule of alkali lignin must contain either fourteen methoxyl groups or some multiple thereof, and that an approximate minimum value for the molecular weight can be obtained from equation (1) by substituting  $n=14$ .

Hence

$$M = \frac{14 \times 31.03}{0.247} = 1760 \dots\dots\dots (16)$$

On the assumption that there are two acidic groups this is in reasonable agreement with the equivalent weight of 863, obtained by potentiometric titration.

Combustion of alkali lignin yielded the figures 62.7 per cent. carbon and 5.9 per cent. hydrogen, corresponding to the formula



On the basis of fourteen methoxyl groups this yields the formula  $\text{C}_{91.4}\text{H}_{102.3}\text{O}_{34.3}$ .

However, the formula that appears to agree best with the data obtained from all derivatives is  $C_{12}H_{14}O_4$ , and this is tentatively accepted as a working hypothesis.

In this case the molecule contains eight hydroxyl groups, six of which react with diamine, the remaining two being presumably anhydric. To determine the nature of these anhydric hydroxyls alkyl lignol was treated with propylaluminumchloride in pyridine. The formation of a secondary derivative showed the presence of one primary alcohol. Alkyl lignol was then treated with benzoyl chloride in pyridine, and it is apparent without hesitation, in pyridine, results were obtained similar to those with acetic anhydride in pyridine, an anhydric derivative being formed, but in pyridine alkyl lignol is a secondary derivative resulted, two hydroxyl groups are reacting. The second anhydric hydroxyl is therefore presumably tertiary, a suggestion supported by the fact that there is one hydroxyl which fails to react with diazotized aniline.

The other hydroxyl which was not benzoylated in the Schotten-Baumann reaction is considered to be present in a tertiary group. The suggestion that the other alkyl lignol is treated with diamine, the resulting hexamethyl alkyl lignol is readily hydrolyzed by cold aqueous alkali to a pentamethyl compound. It would now remain for the fact that octamethyl alkyl lignol is hydrolyzed by heating water to a heptamethyl derivative, in the suggestion that one of the methyl groups is present in a cyclic anhydride.

The possibility in which of derivatives in which the anhydric group is still free, i.e. the hexamethyl and heptamethyl is untenable in view of the large size of the molecule.

Of the other five anhydric hydroxyls at least one is at least a positive reaction being obtained to the Fe, Cl test for acids. Two such hexamethyl which, while stable to cold aqueous alkali, are hydrolyzed by this reagent on the boiling point. The hydrolysis is the accompanied by a secondary reaction whereby two of the hydroxyl groups are eliminated, oxidation of the product yielding a derivative, the size of whose molecule had weight is two less than that of the other derivatives.

Further hexamethyl yields with diamine, a methyl ester, a molecule is easily hydrolyzed in pyridine. Oxidation of heptamethyl alkyl lignol does not yield the expected nonmethylated compound but hexamethylated alkyl lignol. Similarly, octamethyl alkyl lignol yields pentamethylated alkyl lignol, nonmethylated alkyl lignol yields the hexamethylated compound while hydroxymethylated alkyl lignol yields octamethylated compound alkyl lignol, having already two free hydroxyl groups. In this evidence it is interesting to note that Marshall, Brown, and Wittenberg found that their diene ester which, however, when subjected with diamine and was methylated gave a product whose analysis was by no means constant. The diamine derivative contained numerous methyls and three hydroxyls, but the analytical data for the derivative proved agreed more with the proposed lignol but whose methyls and four methyls than with those for the proposed methoxyls and three methyls.

These data are summarized in Table 1.

TABLE I  
ALKALI LIGNIN AND DERIVATIVES

No.	Derivative	Ratio of Hydroxyl/ Methoxyl/ Acyl		OCH <sub>3</sub> (%)	C (%)	H (%)	Acetyl or Benzoyl (%)
1	Alkali lignin	8:14:9	f.	24.7	62.7	5.9	—
			e.	24.7	63.0	6.0	—
2	1 + Ac <sub>2</sub> O	9:14:8	f.	21.1	62.0	5.7	16.8
			e.	20.8	62.1	5.8	16.5
3	2 + H <sub>2</sub> O	1:14:7	f.	21.3	62.1	5.6	14.8
			e.	21.2	62.2	5.8	14.7
4	1 + C <sub>6</sub> H <sub>5</sub> COCl in C <sub>6</sub> H <sub>5</sub> N	0:14:8	f.	16.7	68.7	5.4	32.8
			e.	16.8	68.7	5.3	32.4
5	1 + C <sub>6</sub> H <sub>5</sub> COCl in aq. NaOH	2:14:6	f.	18.2	67.4	5.4	26.2
			e.	18.3	67.7	5.4	26.5
6	1 + CH <sub>2</sub> N <sub>2</sub> .. ..	2:20:0	f.	33.7	63.8	6.2	—
			e.	33.8	64.1	6.4	—
7	6 + cold alc. KOH	3:19:9	f.	32.5	63.7	6.3	—
			e.	32.3	63.8	6.3	—
8	6 + cold alc. NaOH	3:17:9	f.	29.6	63.9	6.2	—
			e.	29.7	64.2	6.1	—
9	8 + Ac <sub>2</sub> O	9:19:3	f.	29.9	63.7	6.3	6.7
			e.	30.2	63.4	6.2	6.6
10	7 + Ac <sub>2</sub> O	6:18:4	f.	28.0	63.2	6.0	8.7
			e.	28.2	63.2	6.1	8.7
11	8 + Ac <sub>2</sub> O	9:16:4	f.	25.4	63.1	6.0	9.0
			e.	25.6	63.4	6.0	8.9
12	2 + Me <sub>2</sub> SO <sub>4</sub> + NaOH	1:21:6	f.	35.2	64.1	6.5	—
			e.	35.1	64.2	6.4	—
13	12 + Ac <sub>2</sub> O .. ..	0:20:2	f.	32.4	63.8	6.3	4.4
			e.	32.2	63.7	6.3	4.5

The presence of two carbonyl groups is shown by the fact that alkali lignin reacts with *p*-nitrophenylhydrazine to give a di-*p*-nitrophenylhydrazone. On treatment with phenylhydrazine under the same conditions alkali lignin yields a phenylhydrazone-osazone, indicating that the primary alcoholic hydroxyl is adjacent to one of these carbonyls.

That active methylene groups are absent is shown by the fact that alkali lignin does not condense with benzaldehyde.

Summing up, therefore, alkali lignin isolated under these conditions has been found to have a minimum molecular weight of 1760, and the empirical formula  $C_{92}H_{104}O_{34}(1754)$  is in agreement with all the data obtained. The molecule contains fourteen methoxyl and eight hydroxyl groups. Of the latter, six are sufficiently acidic to react with diazomethane, and consist of one carboxyl, at least one enol, the remainder being presumably phenolic. There are two alcoholic hydroxyls, one primary and one tertiary.

## II. EXPERIMENTAL

### *General Procedure for the Isolation and Purification of Derivatives*

The following general procedure was used in each case for the isolation and purification of derivatives of alkali lignin. At the completion of a reaction, the reaction mixture was added dropwise with stirring to ten volumes of anhydrous ether, any precipitate filtered off, and the filtrate added to an equal volume of petroleum ether (30–50° C.). If the product was filterable it was filtered at once and washed with petroleum ether. If not, the mixture was allowed to stand overnight when it separated into a clear liquor and a resinous solid adhering to the walls of the vessel. In either case the solid was dissolved in chloroform to give a 5 per cent. solution, filtered, and added dropwise with stirring to ten volumes of anhydrous ether. A small quantity of insoluble material usually separated out at this stage and was filtered off. The filtrate was then added to an equal volume of petroleum ether (30–50° C.), precipitating the main reaction product which was filtered off, washed with petroleum ether, air-dried, and finally dried over sulphuric acid at 20 mm. Hg. This process was repeated until a product of constant methoxyl content was obtained. Prior to analysis the material was dried in an Abderhalden drying pistol over phosphorus pentoxide at 20 mm. Hg. at room temperature.

#### *(a) Alkali Lignin*

From the dioxan-ether mother liquors obtained in the isolation and purification of thiolignin from 200 g. crude lignin(7) the last traces of thiolignin were removed by concentrating to 800 cc. and precipitation with ether (8 l.). After removal of the precipitate (a mixture of thiolignin and alkali lignin, 32 g.) and working up the filtrate, 38 g. alkali lignin was obtained as a fine light buff powder. Concentration of the mother liquors and precipitation yielded a second crop, darker in colour, of 9 g. This was combined with the 32 g. mixed alkali lignin and thiolignin and refractionated, yielding a further 24 g. thiolignin and 8 g. alkali lignin. After purification from chloroform-ether-petroleum ether the final yield of purified alkali lignin was 40 g.

Found: C, 62.7; H, 5.9;  $OCH_3$ , 24.7%; Eq. Wt. 863.

Calculated for  $C_{92}H_{104}O_{34}(1754)$ : C, 63.0; H, 6.0;  $OCH_3$ , 24.7%; Eq. Wt. 877.

Alkali lignin is a yellowish powder when freshly prepared; on standing the colour changes to a light buff. It has no sharp melting point but commences to soften at 112–114° C. It is soluble in ethyl alcohol, acetone, chloroform, acetic acid, ethyl acetate, dioxan, pyridine, and aqueous alkalis, slightly soluble in ether and benzene, and insoluble in petroleum ether, water, dilute acids, and aqueous sodium bicarbonate.

#### *(b) Hexamethyl Alkali Lignin*

Alkali lignin, (50 g.) in dioxan (500 cc.) was methylated with the diazomethane from 35 g. nitrosomethylurea under the conditions used in the preparation of trimethyl thiolignin(7). After



standing at room temperature for two days the mixture was concentrated to 500 cc. and remethylated under the same conditions. After the fourth methylation there was obtained 27.0 g. hexamethyl alkali lignin with a methoxyl content of 33.7 per cent., a value which was unchanged by a subsequent remethylation.

Found: C, 63.8; H, 6.2;  $\text{OCH}_3$ , 33.7%.

Calculated for  $\text{C}_{98}\text{H}_{116}\text{O}_{34}$  (1838): C, 64.1; H, 6.4;  $\text{OCH}_3$ , 33.8%.

Hexamethyl alkali lignin is a bright yellow powder; on standing the colour fades to a dull yellow. It has no sharp melting point, but commences to decompose at 78° C. It is soluble in ethyl alcohol, acetone, chloroform, acetic acid, ethyl acetate, benzene, pyridine, and dioxan, and insoluble in ether, petroleum ether, water, and dilute acids and alkalis.

(c) *Octoacetyl Alkali Lignin*

Alkali lignin (10 g.) in pyridine (50 cc.) was allowed to react with acetic anhydride (15 cc.) for two days at room temperature. After two purifications, yield 6.3 g. Second crop 1.7 g.

Found: C, 62.0; H, 5.7;  $\text{OCH}_3$ , 21.1;  $\text{COCH}_3$ , 16.8%.

Calculated for  $\text{C}_{108}\text{H}_{120}\text{O}_{42}$  (2090): C, 62.1; H, 5.8;  $\text{OCH}_3$ , 20.8;  $\text{COCH}_3$ , 16.5%.

Octoacetyl alkali lignin is a dull yellow powder which commences to soften at 129° C. It is soluble in ethyl alcohol, acetone, chloroform, acetic acid, ethyl acetate, benzene, pyridine, and dioxan, slightly soluble in ether, and insoluble in petroleum ether, water, and dilute acids and alkalis.

(d) *Attempted Methylation of Octoacetyl Alkali Lignin*

Octoacetyl lignin was treated with diazomethane under the conditions used in the preparation of hexamethyl alkali lignin.

Found:  $\text{OCH}_3$ , 21.3%.

Calculated for octoacetyl alkali lignin:  $\text{OCH}_3$ , 20.8%.

(e) *Heptacetyl Alkali Lignin*

Octoacetyl alkali lignin (2 g.) in dioxan (40 cc.) was refluxed with water (10 cc.) for one hour. An excess of water was then added and the precipitated heptacetyl compound filtered off, washed with water, and dried *in vacuo* over  $\text{H}_2\text{SO}_4$ . It was purified twice from chloroform-petroleum ether. Yield 1.3 g.

Found: C, 62.1; H, 5.6;  $\text{OCH}_3$ , 21.3;  $\text{COCH}_3$ , 14.8%.

Calculated for  $\text{C}_{106}\text{H}_{118}\text{O}_{41}$  (2048): C, 62.2; H, 5.8;  $\text{OCH}_3$ , 21.2;  $\text{COCH}_3$ , 14.7%.

Heptacetyl alkali lignin is a light cream powder which commences to soften at 137° C. and is soluble in the same solvents as octoacetyl alkali lignin.

(f) *Octobenzoyl Alkali Lignin*

Alkali lignin (4 g.) in pyridine (40 cc.) was reacted with benzoyl chloride (9 cc.) for two days at room temperature. The mixture was then poured onto crushed ice and allowed to stand overnight in the ice chest. The clear liquor was decanted off and the product, a black tar, dissolved in chloroform (80 cc.). The chloroform solution was successively extracted with saturated aqueous sodium bicarbonate, water, dilute sulphuric acid, and finally water. The solution was then dried by distillation, the volume adjusted to 80 cc., and worked up in the usual way. Yield 2.8 g. Second crop, 0.6 g. The product was purified once more from chloroform-ether-petroleum ether.

Found: C, 68.7; H, 5.4;  $\text{OCH}_3$ , 16.7;  $\text{COC}_6\text{H}_5$ , 32.8%.

Calculated for  $\text{C}_{148}\text{H}_{136}\text{O}_{42}$  (2586): C, 68.7; H, 5.3;  $\text{OCH}_3$ , 16.8;  $\text{COC}_6\text{H}_5$ , 32.4%.

Octobenzoyl alkali lignin is a light buff powder which commences to soften at 156° C. It is soluble in acetone, chloroform, acetic acid, ethyl acetate, benzene, pyridine, and dioxan, and insoluble in ethyl alcohol, ether, petroleum ether, water, and dilute acids and alkalis.

*(g) Hexabenzoyl Alkali Lignin*

Alkali lignin (3 g.) was dissolved in 10 per cent. aqueous sodium hydroxide (30 cc.) and benzoyl chloride (9 cc.) added. After shaking vigorously for a few minutes the reaction product separated out as a brown tarry mass. After standing overnight the product was filtered, washed well with water, and purified once from dioxan-water and once from chloroform-ether-petroleum ether. Yield 2.4 g.

Found: C, 67.4; H, 5.4;  $\text{OCH}_3$ , 18.2;  $\text{COC}_6\text{H}_5$ , 26.2%.

Calculated for  $\text{C}_{134}\text{H}_{136}\text{O}_{40}$ (2378): C, 67.7; H, 5.4;  $\text{OCH}_3$ , 18.3;  $\text{COC}_6\text{H}_5$ , 26.5%.

Hexabenzoyl alkali lignin is a light buff powder which commences to soften at 141° C. It is soluble in the same solvents as octobenzoyl alkali lignin.

*(h) Heptamethyl Alkali Lignin*

Octoacetyl alkali lignin (6 g.) in acetone (80 cc.) was methylated with dimethyl sulphate (60 cc.) and 30 per cent. aqueous sodium hydroxide (160 cc.) under the conditions used in the preparation of hexamethyl thiolignin(7), and the product remethylated to a constant methoxyl content. Yield, after three methylations, 2.0 g. The product was purified three times from chloroform-ether-petroleum ether. Yield 0.8 g.

Found: C, 64.1; H, 6.5;  $\text{OCH}_3$ , 35.2%.

Calculated for  $\text{C}_{99}\text{H}_{118}\text{O}_{34}$ (1852): C, 64.2; H, 6.4;  $\text{OCH}_3$ , 35.1%.

Heptamethyl alkali lignin is a yellow powder which begins to decompose at 118° C. and is soluble in the same solvents as hexamethyl alkali lignin.

*(i) Pentamethyl Alkali Lignin*

Hexamethyl alkali lignin (5 g.) and 5 per cent. ethyl alcoholic potassium hydroxide (30 cc.) were reacted at room temperature for sixteen hours. During this period the hexamethyl derivative dissolved to give a dark brown solution. This was filtered, poured into excess water, and made acid to Congo with dilute sulphuric acid. The precipitated pentamethyl alkali lignin was filtered off, washed with water, and dried *in vacuo*. Yield 4.0 g. After two purifications from chloroform-ether-petroleum ether, yield 2.0 g.

Found: C, 63.7; H, 6.3;  $\text{OCH}_3$ , 32.5%.

Calculated for  $\text{C}_{97}\text{H}_{114}\text{O}_{34}$ (1824): C, 63.4; H, 6.3;  $\text{OCH}_3$ , 32.3%.

Pentamethyl alkali lignin is a yellow powder starting to decompose at 109° C. It is soluble in the same solvents as hexamethyl alkali lignin except that it is soluble in dilute aqueous sodium hydroxide.

*(j) Anhydrotrimethyl Alkali Lignin*

Hexamethyl alkali lignin (2 g.) was refluxed for one hour with 4 per cent. sodium hydroxide in 50 per cent. isopropyl alcohol (20 cc.) and the product isolated as in the previous experiment. Yield 1.7 g. It was purified once from chloroform-petroleum ether.

Found: C, 63.9; H, 6.2;  $\text{OCH}_3$ , 29.6%.

Calculated for  $\text{C}_{95}\text{H}_{108}\text{O}_{33}$ (1778): C, 64.2; H, 6.1;  $\text{OCH}_3$ , 29.7%.

Anhydrotrimethyl alkali lignin is a yellowish buff powder which commences to decompose at 121° C. It is soluble in ethyl alcohol, acetone, chloroform, acetic acid, ethyl acetate, pyridine, dioxan, and aqueous alkali, slightly soluble in benzene, and insoluble in ether, petroleum ether, water, and dilute acids.

*(k) Trityl Alkali Lignin*

Alkali lignin (3 g.) in pyridine (30 cc.) was reacted with triphenylchloromethane (5 g.) at room temperature for forty-eight hours. The mixture was then poured on to crushed ice, yielding a tarry mass which was purified twice from chloroform-ether-petroleum ether. Yield 1.3 g.

Found: C, 66.6; H, 6.0;  $\text{OCH}_3$ , 21.4%.

Calculated for  $\text{C}_{111}\text{H}_{118}\text{O}_{34}$ (1996): C, 66.8; H, 6.0;  $\text{OCH}_3$ , 21.8%.

Trityl alkali lignin is a grey powder decomposing at 70° C. It is soluble in ethyl alcohol, acetone, chloroform, acetic acid, ethyl acetate, dioxan, pyridine, and aqueous sodium hydroxide, and insoluble in ether, petroleum ether, benzene, water, and dilute acids.

(l) *Hexamethyldiacetyl Alkali Lignin*

Heptamethyl alkali lignin (0.7 g.) in pyridine (10 cc.) was reacted with acetic anhydride (1.5 cc.) for forty-eight hours at room temperature. The product was isolated and purified as in the case of octoacetyl alkali lignin. Yield (purified) 0.5 g.

Found: C, 63.8; H, 6.3;  $\text{OCH}_3$ , 32.4;  $\text{COCH}_3$ , 4.4%.

Calculated for  $\text{C}_{108}\text{H}_{120}\text{O}_{36}$ (1922): C, 63.7; H, 6.3;  $\text{OCH}_3$ , 32.2;  $\text{COCH}_3$ , 4.5%.

Hexamethyldiacetyl alkali lignin is a yellow powder which begins to decompose at 124° C. It is soluble in acetone, chloroform, acetic acid, ethyl acetate, benzene, dioxan, and pyridine, slightly soluble in ethyl alcohol, and insoluble in ether, petroleum ether, water, and dilute acids and alkalis.

(m) *Pentamethyltriacetyl Alkali Lignin*

Hexamethyl alkali lignin (10 g.) in pyridine (50 cc.) was reacted with acetic anhydride (15 cc.) and the product worked up as for the octoacetyl derivative. Yield, after five purifications from chloroform-petroleum ether, 4.8 g.

Found: C, 63.7; H, 6.3;  $\text{OCH}_3$ , 29.9;  $\text{COCH}_3$ , 6.7%.

Calculated for  $\text{C}_{108}\text{H}_{120}\text{O}_{37}$ (1950): C, 63.4; H, 6.2;  $\text{OCH}_3$ , 30.2;  $\text{COCH}_3$ , 6.6%.

Pentamethyltriacetyl alkali lignin is a yellow powder which begins to decompose at 115° C. and is soluble in the same solvents as octoacetyl alkali lignin.

(n) *Tetramethyltetracetyl Alkali Lignin*

Pentamethyl alkali lignin (1.6 g.) in pyridine (16 cc.) was reacted with acetic anhydride (3 cc.), and the product worked up as for the octoacetyl derivative. Yield (purified) 1.2 g.

Found: C, 63.2; H, 6.0;  $\text{OCH}_3$ , 28.0;  $\text{COCH}_3$ , 8.7%.

Calculated for  $\text{C}_{104}\text{H}_{120}\text{O}_{38}$ (1978): C, 63.2; H, 6.1;  $\text{OCH}_3$ , 28.2;  $\text{COCH}_3$ , 8.7%.

Tetramethyltetracetyl alkali lignin is a yellow powder which begins to decompose at 131° C. and is soluble in the same solvents as hexamethyldiacetyl alkali lignin.

(o) *Anhydrodimethyltetracetyl Alkali Lignin*

Anhydrotrimethyl alkali lignin (1 g.) in pyridine (5 cc.) was reacted with acetic anhydride (1.5 cc.), and the product worked up as for octoacetyl alkali lignin. Yield (purified) 0.7 g.

Found: C, 63.1; H, 6.0;  $\text{OCH}_3$ , 25.4;  $\text{COCH}_3$ , 9.0%.

Calculated for  $\text{C}_{102}\text{H}_{114}\text{O}_{37}$ (1932): C, 63.4; H, 6.0;  $\text{OCH}_3$ , 25.6;  $\text{COCH}_3$ , 8.9%.

Anhydrodimethyltetracetyl alkali lignin is a yellow powder which begins to decompose at 106° C. and is soluble in the same solvents as hexamethyldiacetyl alkali lignin.

(p) *Alkali Lignin Phenylhydrazono-Osazone*

Alkali lignin (5 g.) in glacial acetic acid (100 cc.) was refluxed for three hours with redistilled phenylhydrazine (6 cc.), yielding 0.7 g. of a chloroform-insoluble tar and 4.2 g. of a chloroform-soluble fraction. Fractionation of the latter yielded 1.2 g. of a chloroform-ether insoluble fraction and 1.85 g. of a chloroform-ether soluble fraction.

Found: (1) chloroform insoluble fraction:  $\text{OCH}_3$ , 14.0; N, 5.4%.

(2) chloroform-ether insoluble fraction:  $\text{OCH}_3$ , 18.9; N, 4.4%.

(3) chloroform-ether soluble fraction: C, 65.7; H, 6.0; N, 4.0;  $\text{OCH}_3$ , 21.5%.

Calculated for  $\text{C}_{110}\text{H}_{120}\text{N}_6\text{O}_{31}$ (2022): C, 65.3; H, 6.0; N, 4.2;  $\text{OCH}_3$ , 21.5%.

Alkali lignin phenylhydrazono-osazone is an orange-yellow powder when first prepared; the colour changes on standing to a light brown. It begins to decompose at 117° C. It is soluble in ethyl alcohol, chloroform, acetic acid, ethyl acetate, pyridine, dioxan, and dilute sodium hydroxide, slightly soluble in ether and benzene, and insoluble in petroleum ether, water, and dilute acids.

(g) *Alkali Lignin Di-p-Nitrophenylhydrazone*

This preparation was carried out under conditions similar to those used in the preceding experiment. From 5 g. alkali lignin 4.2 g. of crude product was obtained, yielding 2.0 g. of purified chloroform-ether soluble fraction.

Found: C, 61.6; H, 5.9; N, 4.1;  $\text{OCH}_3$ , 21.4%.

Calculated for  $\text{C}_{104}\text{H}_{114}\text{N}_6\text{O}_{36}$ (2024): C, 61.7; H, 5.7; N, 4.2;  $\text{OCH}_3$ , 21.4%.

(r) *Attempted Condensation of Alkali Lignin with Benzaldehyde*

Alkali lignin (4 g.) in 10 per cent. aqueous sodium hydroxide (40 cc.) was reacted with benzaldehyde (2.5 g.) under the conditions used to prepare benzalthiolignin(8). Yield 3.3 g. After two purifications from chloroform-petroleum ether, yield 2.2 g.

Found:  $\text{OCH}_3$ , 24.5%.

Calculated for unchanged alkali lignin:  $\text{OCH}_3$ , 24.7%.

### III. ACKNOWLEDGMENTS

The author is indebted to Messrs. Australian Paper Manufacturers Limited for permission to publish this work: to Dr. F. N. Lahey of the University of Melbourne, co-author of an earlier paper in this series, for his continued interest and advice: and to Miss Betty Hickox and Mr. N. Lottkowitz for semi-micro- and microanalyses respectively.

### IV. REFERENCES

- (1) PHILLIPS, M.—Wise's "Wood Chemistry." 1st Ed., p. 286. (Reinhold: New York, 1944.)
- (2) BROOKBANK, E. B., BRAUNS, F. E., LEWIS, H. F., and BUCHANAN, M. A.—*Paper Tr. J.* **116** (13): 27 (1943).
- (3) POWTER, N. B.—*Paper Ind. Paper World* **28**: 1744 (1947).
- (4) PLUNGIAN, M.—*Industr. Engng. Chem.* **32**: 1399 (1940).
- (5) TOMLINSON, G. H., and TOMLINSON, G. H. II.—U.S. Pat. 2,406,867 (Sept. 3, 1946).
- (6) AHLM, C. E.—*Paper Tr. J.* **113** (13): 115 (1941).
- (7) LAHEY, F. N., and MEREWETHER, J. W. T.—*Aust. J. Sci. Res. A* **1**: 112 (1948).
- (8) MEREWETHER, J. W. T.—*Ibid.* **1**: 241 (1948).
- (9) MARSHALL, H. B., BRAUNS, F. E., and HIBBERT, H.—*Canad. J. Res. B* **13**: 103 (1935).



# THE CHEMICAL CONSTITUENTS OF AUSTRALIAN *FLINDERSIA* SPECIES\*

## I. COLLININ, 7-GERANOXY-8-METHOXYCOUMARIN

By F. A. L. ANET,† F. R. BLANKS,† and G. K. HUGHES†

[Manuscript received December 17, 1948]

### Summary

*Collinin*, a new coumarin, has been isolated from the bark of *Flindersia collina* Bail. and shown to be 7-geranoxy-8-methoxycoumarin.

## I. INTRODUCTION

The genus *Flindersia* consisting of some twenty-one species, fourteen of which are found in Australia, is placed by Engler and Prantl(1) in the family Rutaceæ, subfamily Flindersioideæ, tribe Flinderseæ and was formerly noted by Brown(2) to be closely related to the tribe Xanthoxyleæ. The Australian species are mostly found in the rain-forest and provide some well-known commercial timbers. The only previous chemical investigation in this genus was carried out by Matthes and Schreiber(3) who claimed the isolation of an alkaloid "Flindersine" of unknown constitution. Because of this report, the availability, and the comparative ease of collection, it was decided systematically to survey the barks of the Australian species for their chemical constituents.

This paper reports the isolation of a coumarin from the bark of *Flindersia collina* Bail. This tree, commonly known as "leopard ash", has a fairly restricted distribution in eastern Australia. The bark is characteristically mottled and patchy, the leaves are compound and when immature are notable in that the leaf rhachis is winged.

A prollius extract gave positive alkaloid tests both on freshly collected and oven-dried (140° F.) material‡ (L. J. Webb, personal communication). A petroleum ether extract of the bark on concentration deposited crystals which were easily purified to give a colourless crystalline substance, herein called *collinin*, which analysed for  $C_{20}H_{20}O_4$ . During a methoxyl determination, which showed the presence of one methoxyl group, it was noticed that a sweet

\* This plant material was supplied by Mr. L. J. Webb, C.S.I.R., and is part of the survey on the chemical constituents of the Australian flora being carried out by C.S.I.R. and other institutions.

† Department of Chemistry, University of Sydney.

‡ Two colourless crystalline alkaloids have been isolated in small yield and will be reported in a later communication.

smelling oil was produced. A similar oil was produced by distilling collinin with constant boiling hydrochloric acid and it was shown to be geranyl chloride.

On cooling the residual mother liquors from this experiment a white crystalline product (I) was precipitated which analysed for  $C_9H_4O_2(OH)(OCH_3)$ . (I) was also isolated from a concentrated methanolic extract of the bark by boiling it with dilute hydrochloric acid. This substance was easily mono-acetylated to  $C_9H_4O_2(OCH_3)(OCOCH_3)$  and monomethylated by diazomethane or methyl iodide and potassium carbonate to  $C_9H_4O_2(OCH_3)_2$ . However, when its yellow solution in caustic soda was treated with dimethyl sulphate three products were isolated all of which were carboxylic acids. Two were isomeric, having the composition  $C_8H_4(OCH_3)_3COOH$ , and the third analysed for  $C_8H_4(OH)(OCH_3)_2COOH$ . These reactions and products suggested that (I) was an hydroxymethoxycoumarin and it was finally identified as the known daphnetin 8-methyl ether (7-hydroxy-8-methoxycoumarin).\*

The dimethoxy acid (II) (m.p.  $210-11^\circ C.$ ) on exposure to ultraviolet light was converted to (I) and consequently is 2-hydroxy-3,4-dimethoxycinnamic acid. This has been previously described by Gruber(8) as a yellow solid m.p.  $202^\circ C.$ , but this product must have been impure as on repeating his experiment there was obtained a colourless acid m.p.  $210-11^\circ C.$  identical with (II).

The two trimethoxy acids (IV) and (III) are most probably the *trans* and impure *cis* forms of 2,3,4-trimethoxycinnamic acid. The low melting one (m.p.  $72^\circ C.$  (III)) was readily converted to the other by heating above its melting point or by repeated recrystallization from water. The melting point of  $72^\circ C.$  is not claimed as the true value of the *cis* form. The higher melting acid has been described by Slotta and Heller(9) and their method of synthesis suggests that it has the *trans* configuration.

The formation of geranyl chloride and daphnetin methyl ether from collinin is understandable since it is well known that allyl ethers are readily split by acids. This leads to the conclusion that collinin is 7-geranoxy-8-methoxycoumarin (V), and this was confirmed by heating (I) and geranyl chloride in the presence of sodium ethoxide. It is interesting to note that an isomer of (V), 5-geranoxy-7-methoxycoumarin, has been recently isolated by Caldwell and Jones(10) from expressed West Indian lime oil.

## II. EXPERIMENTAL

### (a) Isolation of Collinin

The ground bark (1.8 kg.) was continuously extracted with petroleum ether (b.p.  $60-90^\circ C.$ ) for several days, the solution concentrated to 200 ml., and refrigerated. The resulting precipitate

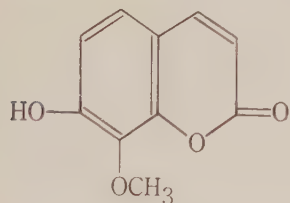
\* Hattori(4) who prepared (I) by methylation and hydrolysis of natural daphnin states the melting point to be  $158^\circ C.$  Wessely and Sturm(5) made it in the same way and recorded a melting point of  $185.5^\circ C.$  Gaudin(6) who obtained it by partial methylation of daphnetin with diazomethane gave melting point as  $185^\circ C.$  Finally Spath and Schmid(7) synthesized it from 2,4-dihydroxy-3-methoxybenzaldehyde and reported the melting point as  $161^\circ C.$  In view of this discordance the undoubted synthesis of Spath and Schmid was repeated and the product had a melting point of  $158^\circ C.$  (uncorrected) undepressed by the natural phenol (I).

was dissolved in ether, the solution washed with sodium hydroxide and hydrochloric acid, the solvent removed, and the residue recrystallized from petroleum ether. Yield 20 g., m.p. 67–68° C.

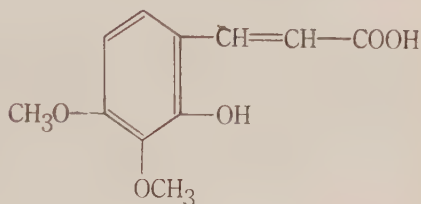
Found: C, 72.5; H, 7.61; MeO, 9.15%.

Calculated for  $C_{20}H_{20}O_4$ : C, 73.1; H, 7.36; 1MeO, 9.45%.

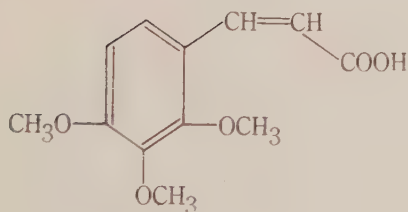
It is insoluble in hot and cold water but dissolves in all the common organic solvents and warm sodium hydroxide to give a yellow solution.



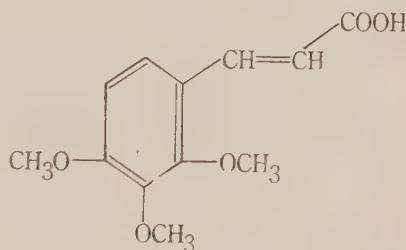
(I)



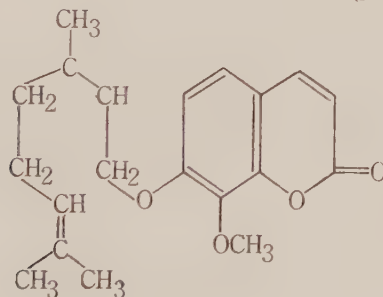
(II)



(III)



(IV)



(V)

#### (b) Hydrolysis of Collinin

Collinin (2 g.) was heated with constant boiling hydrochloric acid (25 ml.) in a distillation flask fitted with a water condenser until it had all dissolved and no more oily drops were condensed.

The residual liquor was diluted with water, cooled, the precipitate collected, and recrystallized from water. Yield 1 g., m.p. 158° C. identical with daphnetin 8-methyl ether as prepared below. From the distillate an oil b.p. 184–7° C. (0.5 g.) was recovered. This was dissolved in alcohol (2 ml.), thiourea (0.5 g.) added, boiled for a few minutes, and a little saturated alcoholic picric acid added with cooling. The precipitate after recrystallization from alcohol had m.p. 137–8° C.

Found: N, 15.9%.

Calculated for  $C_{17}H_{23}O_7N_2S$ : N, 15.8%.

This was identical with a specimen of *S*-geranylisothioureia picrate prepared from authentic geranyl chloride in the same way.

*(c) Isolation of Daphnetin 8-Methyl Ether*

The ground bark was extracted with alcohol continuously for several days. After concentrating the deep brown solution, the thick syrup was boiled with dilute hydrochloric acid. The hot aqueous layer was decanted and on cooling deposited large slightly pink crystals which, after recrystallization from water (charcoal), had m.p. 158° C.

Found: C, 62.2; H, 4.25; MeO, 15.6%.

Calculated for  $C_{16}H_8O_4$ : C, 62.5; H, 4.19; 1MeO, 16.25%.

*(d) Synthesis of Daphnetin 8-Methyl Ether*

The method was essentially that of Späth and Schmid (loc. cit.) except that the acetate was isolated and purified before hydrolysis. This had m.p. 135–6° C. undepressed on admixture with the acetate obtained from the natural product. The acetate was hydrolysed by refluxing with constant-boiling hydrochloric acid for five minutes to give after recrystallization a product m.p. 158° C.

*(e) Daphnetin 8-Methyl Ether Acetate (ex. F. collina)*

Prepared in the usual way and recrystallized from water.

Found: C, 61.3; H, 4.28;  $CH_3CO$ , 18.5%.

Calculated for  $C_{12}H_{10}O_5$ : C, 61.5; H, 4.30;  $1CH_3CO$ , 18.4%.

*(f) Daphnetin Dimethyl Ether (ex. F. collina)*

(i) (I) (1 g.) was dissolved in absolute alcohol, excess ethereal diazomethane added, and left overnight. The product, after removal of solvent, crystallized from water as colourless needles m.p. 118–9° C. (lit. 119–21° C.).

(ii) (I) (2 g.), acetone (50 ml.), methyl iodide (10 ml.), and potassium carbonate (6 g.) were refluxed for 5 hours and the product, isolated in the usual way, was identical with the above.

Found: C, 63.35; H, 4.90; MeO, 28.9%.

Calculated for  $C_{11}H_{10}O_4$ : C, 64.0; H, 4.89; 2MeO, 30.1%.

*(g) Action of (I) with Dimethyl Sulphate and Caustic Soda*

(I) (2 g.), methanol (16 ml.), and sodium hydroxide solution (30 ml. 20%) were warmed to 50° C. and dimethyl sulphate (12 ml.) gradually added with stirring. After cooling, the solution was made just acid and the precipitate (A) collected, m.p. 70–86° C. On standing for some days the mother liquors deposited a further crop (B), m.p. 157–160° C. (A) was partly soluble in cold benzene, the residue on recrystallization from water gave 2-hydroxy-3,4-dimethoxycinnamic acid, m.p. 209–10° C.

Found: C, 58.3; H, 5.52; MeO, 27.5%.

Calculated for  $C_{11}H_{12}O_5$ : C, 58.9; H, 5.40; 2MeO, 27.5%.

The benzene soluble portion after crystallizing from water gave a product m.p. 72° C. which after heating at 100° C., cooling, and reheating, melted at 168–9° C.

On recrystallizing from water (B) gave *trans*-2,3,4-trimethoxycinnamic acid, m.p. 168–9° C.

Found: C, 59.5; H, 5.92; MeO, 37.5%.

Calculated for  $C_{12}H_{14}O_5$ : C, 60.5; H, 5.92; 3MeO, 39.1%.

*(h) Synthesis of Collinin*

(I) (1 g.), geranyl chloride (0.9 g.), sodium ethoxide (from 0.12 g. of sodium), and alcohol (10 ml.) were refluxed for 8 hours. The product, isolated in the usual way, on recrystallization from petroleum ether had m.p. 67° C. undepressed by natural collinin.

## III. ACKNOWLEDGMENTS

The authors wish to thank Mr. L. J. Webb, C.S.I.R., for bringing to our notice the possibilities for chemical investigation in the *Flindersias*, the collection



of bark, and for the botanical description. We are also indebted to Miss J. E. Fildes, Microanalyst, University of Sydney, for all the analyses.

#### IV. REFERENCES

- (1) ENGLER, A., and PRANTL, K.—“Die natürlichen Pflanzen-familien.” (Engelmann : Leipzig, 1931.)
- (2) BAILEY, F. M.—“Queensland Flora.” Vol. I, p. 238. (Govt. Printer : Brisbane.)
- (3) MATTHES, H., and SCHREIBER, E.—*Ber. dtsh. pharm. Ges.* **24** : 385 (1914).
- (4) HATTORI, S.—*J. Pharm. Soc. Japan* **50** : 82 (1930).
- (5) WESSELY, F., and STURM, K.—*Ber. dtsh. chem. Ges.* **63** : 193 (1941).
- (6) GANDINI, A.—*Gazz. Chim. Ital.* **70** : 611 (1940).
- (7) SPÄTH, F., and SCHMID, H.—*Ber. dtsh. chem. Ges.* **74** : 193 (1941).
- (8) GRUBER, W.—*Mh. Chem.* **75** : 14 (1944).
- (9) SLOTTA, K., and HELLER, H.—*Ber. dtsh. chem. Ges.* **63** : 3029 (1930).
- (10) CALDWELL, A., and JONES, E.—*J. Chem. Soc.* **1945** : 540 (1945).

# THE OCCURRENCE OF $\beta$ -CAROTENE AND "PHYTOFLUENE" IN THE WOOD OF *ACACIA ACUMINATA*

By P. STANLEY\* and V. M. TRIKOJUS\*

[Manuscript received November 5, 1948]

## Summary

The separation from extracts of *Acacia acuminata*, by chromatographic procedures, of  $\beta$ -carotene and a colourless fluorescent oil of similar adsorption properties is described.

On the basis of its absorption spectrum and other properties the fluorescent oil is considered to be closely related to and possibly identical with the colourless carotenoid, phytofluene.

## I. INTRODUCTION

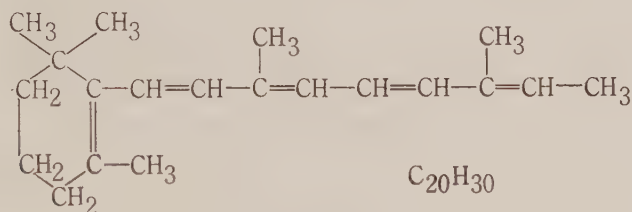
*Acacia acuminata* Benth. is a tree indigenous to the temperate woodlands of Western Australia, where it attains a height of 20-30 or even 40 ft. The timber is rich, reddish-brown in colour, very hard, with interlocked grain and very durable in the ground. The tree is known colloquially as "jam wood", owing to the fact that when freshly cut there is a strong odour, suggestive of raspberry jam. In a preliminary communication Trikojus and Drummond (1) reported the results of an examination of the oil obtained from the wood by percolation with petroleum ether. They isolated from the oil, by chromatography,  $\beta$ -carotene, which was obtained in crystalline form, and they observed also that highly fluorescent substances were present. As far as the present authors are aware this still remains the only reported occurrence of  $\beta$ -carotene in a timber extract. However, several other members of the genus *Acacia*, namely, *A. harpophylla* F. v. M., *A. pendula* A. Cunn., *A. cambagei* R. T. Baker, and *A. rhodogydon* Maiden, provide timbers as dense or denser than that of *Acacia acuminata* and, in addition, each of these timbers is reported to have, when freshly cut, a similar though fainter odour.

The discovery by Zechmeister and Sandoval (2, 3) in a range of plant extracts of a fluorescent hydrocarbon, phytofluene, with characteristic absorption spectrum, led us to investigate more fully the substances present in petroleum ether extracts of the wood of *Acacia acuminata*. A colourless fluorescent oil has been shown to be present in the wood; it has been purified and certain of its properties have been defined. Its adsorption characteristics, absorption spectrum, and specific extinction, and the results of the Carr Price test and microanalysis agree well with those indicated by Zechmeister and Sandoval (3) for phytofluene. Consequently we have, for convenience in the text, given this name also to our substance although we do not claim to have established its identity beyond all doubt.

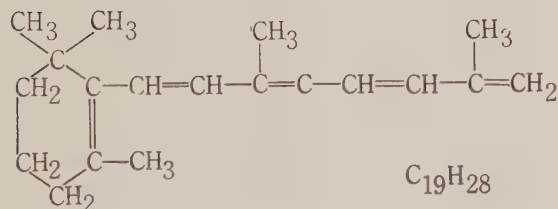
\* Department of Biochemistry, University of Melbourne.

The occurrence of phytofluene in the wood of this tree, along with  $\beta$ -carotene, is consistent with the observations by Zechmeister and Sandoval(3), firstly, that phytofluene occurs in those parts of a plant free from chlorophyll, and secondly, that it is usually present in association with carotenoid pigments. The concentration of phytofluene in the wood, 26–28 mg./kg. whole wood meal (containing 25.0 per cent. water), indicates that this material may act as a useful source of phytofluene for further chemical investigations. In another sample of the same log the concentration of  $\beta$ -carotene was estimated spectrophotometrically to be of the order of 200 mg./kg. whole wood meal (containing 23.1 per cent. water), but only a minor portion of the pigment could be isolated in crystalline form(1).

Several compounds recently described, in addition to those mentioned by Zechmeister and Sandoval(3), possess absorption spectra closely related to that of phytofluene. Thus axerophthene (I), the parent hydrocarbon of vitamin A, has, in ethanol, absorption maxima at 331, 346, and 364 m $\mu$ (4), while its lower



(I)



(II)

homologue (II), also in ethanol, has its maxima at 330, 348, and 367 m $\mu$ , with  $E_1^{1\%} = 1800$  at 348 m $\mu$ (5).\* In the same solvent, phytofluene (isolated from *Acacia acuminata*) has absorption maxima located at 331.5, 348, and 368 m $\mu$ . While (I) has been reported as exhibiting vitamin A activity(4), neither (II), according to Shantz(5), nor phytofluene(3) possess this property. (I) and (II) are stated to be pale orange oils, while phytofluene can be obtained as a colourless oil. Although the resemblance of the absorption spectra of these compounds suggests that there is a common element in their structures, deductions as to the constitution of phytofluene should await further chemical investigation on this substance.

\* See also Karrer, P., and Davis, P. C.—*Helv. Chim. Acta* **31**: 1607 (1948).

## II. EXPERIMENTAL

## (a) Materials

*Timber*.—A freshly cut log of *Acacia acuminata*, about 6 ft. long and 8 in. diameter, was obtained from its natural habitat (Western Australia). The ends of the log were sealed with paint and it was kept in a cool room and sectioned as required for extraction purposes.

*Solvents*.—All solvents were purified before use. Commercial petroleum ether was extracted for several days with concentrated  $\text{H}_2\text{SO}_4$ , containing about 40 ml. of 20 per cent. oleum per l., followed by a saturated solution of  $\text{KMnO}_4$  in 10 per cent.  $\text{NaOH}$  for 24 hr. It was then dried over  $\text{NaOH}$  flakes and distilled, the fractions b.p. 40–55° C., 55–65° C. being collected separately. Ethanol and methanol were refluxed over amalgamated Al-turnings and distilled.

*Adsorbents*.—A number of adsorbents was tested, but finally only alumina (B.D.H. "For Chromatography"), activated at 150° C. for 3–5 hr. and partly de-activated with 0.7 g. 100 g. water, was employed. This was regenerated after use by washing with commercial ethanol, until the filtrate was colourless, and oxidation of residual organic matter by heating at 170° C. for several days, the pure white of the original material being thus regained. No difference between chromatograms on fresh and reclaimed alumina could be detected (6, 7).

## (b) Extraction and Separation of Pigments and Phytofluene

For extraction, sections of the log of *Acacia acuminata* were sawn into disks of 1 cm. thickness which were split with a chisel and the pieces ground in an electric Wiley mill to pass a 1 mm. sieve; alternatively, the material was converted into sawdust, using a nest of circular saws. The meal was stored at room temperature in closed bottles. No change in the phytofluene content was observed during 2 months' storage under these conditions.

Preliminary experiments showed that extraction of the meal in a soxhlet with low boiling solvents did not cause destruction of the pigments or phytofluene. It is also interesting to note that xanthophylls appeared to be absent, except possibly in traces, as determined by adsorption of the crude extract on dicalcium phosphate ( $\text{CaHPO}_4$ ; Merck "Pure")(8).

Batches of 250 g. were extracted in a glass soxhlet with petroleum ether (b.p. 40–55° C.) the soxhlet being allowed to operate for 1 hr. This extraction and all subsequent operations were carried out in a darkroom with low intensity artificial white or orange-red light. The crude extract was not suitable for the spectrophotometric estimation of phytofluene at 348 m $\mu$  (Beckman Spectrophotometer, Model DU) as there was a considerable amount of irrelevant absorption at that wavelength. The extract was therefore chromatographed on a column of alumina (1.3 cm.  $\times$  40 cm.), the chromatogram being developed with petroleum ether (b.p. 55–65° C.). Preliminary concentration of the extract was not found to be necessary. Development was continued until about 1 l. of filtrate was obtained. A typical chromatogram showed the following zones:

- 5 mm. pale yellow
- 30 mm. orange-brown ( $\beta$ -carotene and oxidation products)
- 180 mm. orange (mostly  $\beta$ -carotene)
- 6 mm. pale yellow (possibly *neo*- $\beta$ -carotenes).

There was some variation in the length of the individual zones, but they were always present in the order mentioned.

When such a column was observed in ultraviolet light, the highest zone of 5 mm. was found to contain substances with pale blue fluorescence, but without fine absorption spectra. Below the lower pale yellow zone (6 mm.), and overlapping with it, a bright, greenish-yellow fluorescence was observed. This zone (50 mm.) was found to contain the phytofluene, for the estimation of which the column was cut so as to include all of the fluorescent material of the zone, irrespective of pigments also present.

## (c) Phytofluene

*Estimation*.—Phytofluene was eluted by leaving the adsorbent overnight in contact with 50 ml. of petroleum ether (b.p. 55–65° C.) and 20 ml. of methanol at 0° C. under  $\text{N}_2$ . The adsorbent



was then washed with further small amounts of petroleum ether, until no greenish-yellow fluorescence persisted on the adsorbent. The combined eluates were washed free of methanol with water (5-6 washings), dried over  $\text{Na}_2\text{SO}_4$ , diluted to 250 ml. and 10 ml. of this concentrated solution diluted to 100 ml. Phytofluene was estimated spectrophotometrically in the diluted solution at its absorption maximum located within the range 347 to 348  $\text{m}\mu$ , using the value  $E_{1\text{ cm.}}^{1\%} = 1200(2, 3)$ .

To estimate the losses of phytofluene occurring during chromatography (due to incomplete elution and other factors), the concentrated phytofluene solution was rechromatographed and phytofluene eluted from this second and from subsequent columns was estimated. The results are given in Table 1. For comparison, figures obtained by Zechmeister and Sandoval(2) during the purification of phytofluene from tomato paste are included. The losses of phytofluene per chromatographic run are seen to be roughly parallel.

TABLE 1

PHYTOFLUENE CONTENT OF *Acacia acuminata* ESTIMATED SPECTROPHOTOMETRICALLY

Experiment	Phytofluene (mg./kg. Wood Meal, containing 25.0% $\text{H}_2\text{O}$ ) after $n$ Chromatographic Runs on $\text{Al}_2\text{O}_3$ where :				
	$n=1$	2	3	4	5
1	27.3				
2	20.9				
3	25.6				
4	25.4				
5	23.8	12.7	10.0		
6	31.4				
7	28.9	20.3	16.8	13.1	
8	29.2	20.0	18.3	14.6	
9	25.3				
Z. & S.*	20.2	16.7	15.1	11.2	9.5

\* Zechmeister and Sandoval(3); values as mg./kg. tomato paste. Adsorbent for 1st chromatographic run  $\text{Ca}(\text{OH})_2$ , for the others mixtures of alumina and  $\text{Ca}(\text{OH})_2$ .

*Purification.*—The phytofluene solutions used for the estimations were collected and stored in the dark under  $\text{N}_2$  at  $0^\circ\text{C}$ . Phytofluene extracts were also prepared from wood meal extracted in a metal soxhlet in 1 kg. portions and chromatographed once. The combined petroleum ether extracts were treated overnight with a fresh solution of  $\text{NaOEt}$  in ethanol at room temperature in the dark and under  $\text{N}_2$ . The solution was washed free of ethanol, dried, and chromatographed several times on alumina, until little coloured material remained. The last column was cut quickly in white light so as to exclude all pigments, the wholly white zone on elution yielding phytofluene free of pigments. The eluate, washed with water and dried ( $\text{Na}_2\text{SO}_4$ ), was made up to a calculated concentration of 8-12 mg./100 ml. in petroleum ether (b.p.  $40-55^\circ\text{C}$ ). Of 50 ml. of such a solution in a standard flask, 5 ml. was diluted to 100 ml. for spectrophotometric determination of the optical density, the balance being evaporated in a current of  $\text{N}_2$  or  $\text{CO}_2$  in one of two 20 ml. conical flasks with ground necks. Last traces of solvent were removed by evacuation in a desiccator and the colourless residual oil weighed on a microbalance.

Found: for specimen (a),  $E_{1\text{ cm.}}^{1\%} \times 10^{-3}$  at 348  $\text{m}\mu = 1.26$ ; for specimen (b),  $E_{1\text{ cm.}}^{1\%} \times 10^{-3}$  at 331  $\text{m}\mu = 0.87$ , 338  $\text{m}\mu = 0.73$ , 348  $\text{m}\mu = 1.33$ , 358  $\text{m}\mu = 0.66$ , 367  $\text{m}\mu = 1.18$ . These  $E_{1\text{ cm.}}^{1\%}$ .

values are slightly higher than those recorded by Zechmeister and Sandoval(3). The curve of the absorption spectrum of phytofluene as depicted in Figure 1 follows very closely that given by the authors just mentioned.

After evaporation of 5 ml. portions of the dilute solution of specimen (b) and dissolution in *n*-hexane, the  $E_{1\text{ cm}}^{1\%}$  values for the hexane solution were found to be identical with those obtained using petroleum ether, within the limits of experimental error.

*Carr-Price Test.*—Addition of a saturated solution of  $\text{SbCl}_3$  in chloroform (stabilized with a little acetic anhydride) to a chloroform solution of phytofluene caused the immediate development of a blue colour, which changed rapidly through permanganate red to pale brown(3).

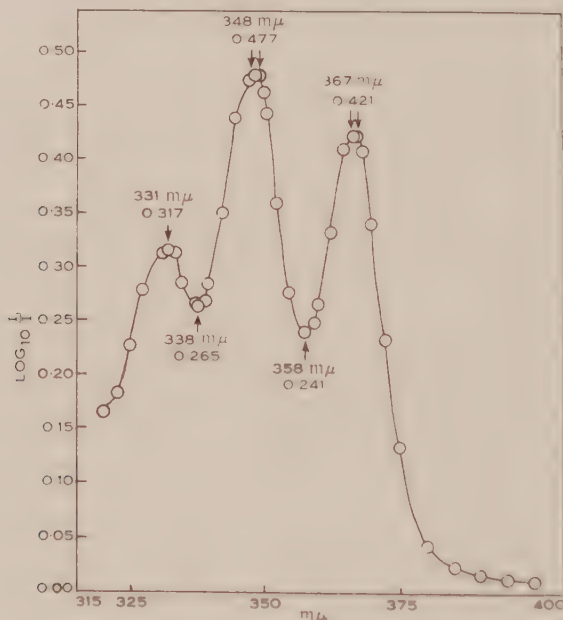


Fig. 1.—Absorption spectrum of phytofluene from *Acacia acuminata*. Solvent, petroleum ether, b.p. 40–55° C. Conc. 0.353 mg./100 ml.

*Microanalysis.*—Analysis was attended by considerable difficulty owing to the instability of the solvent-free material(3). Some of the petroleum ether solution of specimen (a) was concentrated, the last 2 ml. being evaporated in a platinum combustion boat in a current of  $\text{N}_2$ .

Found: C, 84.50; H, 10.95%; corrected for 4.55%  $\text{O}_2$  (by difference), C, 88.53; H, 11.47%.

Calculated for  $\text{C}_{40}\text{H}_{64}$ : C, 88.15; H, 11.85%.

The corrected values are in accord with the molecular formula  $\text{C}_{40}\text{H}_{64.2}$  proposed by Zechmeister and Sandoval(3).

#### (d) $\beta$ -Carotene

*Estimation.*—Up to the cutting of the column, the procedure was the same as for the estimation of phytofluene. The whole of the orange-brown and orange zones was now cut out and eluted as before and the dried petroleum ether solution made up to 250 ml. Of this solution 5 ml. was diluted to 250 ml. and the  $\beta$ -carotene determined spectrophotometrically at the main absorption maximum located within the range 442–444 mμ. The second maximum was located in the range 469–471 mμ. These maxima are, therefore, considerably displaced from those

quoted(9) for pure  $\beta$ -carotene (452 and 478  $m\mu$ ), heat-treated  $\beta$ -carotene (451 and 477-9  $m\mu$ ) or for iodine-isomerized  $\beta$ -carotene (447 and 473  $m\mu$ ). Even complete isomerization could not, therefore, account for the displacement of the maxima. This was attributed to the presence in

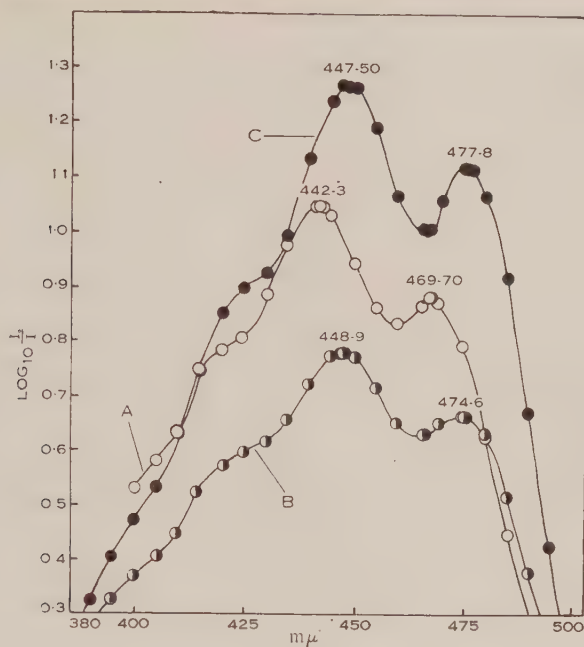


Fig. 2.—Absorption spectra of preparations of  $\beta$ -carotene from *Acacia acuminata*. Curve A. Eluted " $\beta$ -carotene zone" from chromatographed crude extract (petroleum ether, b.p. 55-65° C.). Approx. conc. 0.40 mg./100 ml. Curve B. Crude crystals of  $\beta$ -carotene (hexane). Approx. conc. 0.30 mg./100 ml. Curve C. Recrystallized  $\beta$ -carotene (hexane). Approx. conc. 0.24 mg./100 ml.

the solution of appreciable amounts of an unidentified oily substance, which could not be removed by saponification and traces of which remained even after several chromatographic purifications.

TABLE 2

$\beta$ -CAROTENE CONTENT OF *Acacia acuminata* ESTIMATED SPECTROPHOTOMETRICALLY  
(APPROXIMATE VALUES)

Experiment	1	2	3	4	5
$\beta$ -Carotene (mg./kg. wood meal containing 23.1% H <sub>2</sub> O)	199	200	189	189	195

As can be seen from Figure 2, the shape of the absorption curve of the solutions used for estimating the  $\beta$ -carotene content of the wood (Curve A) differs little from that of the purified pigment from the same source. The absorption maxima, in the latter case, and the general shape of the curve

are also very close to those reported elsewhere for  $\beta$ -carotene(9). However, the values obtained (Table 2), applying the factor  $E_{1\text{ cm.}}^{1\%} = 2600$  given by Devine, Hunter, and Williams(10) for highly purified  $\beta$ -carotene at 453  $m\mu$ , must be regarded as approximate only.

*Purification.*—The combined  $\beta$ -carotene solutions, as used for the estimations, were kept at 0° C. in the dark under  $N_2$  (total content c. 900 mg.). Esters present in the solution were saponified with a solution of NaOEt (from 10 g. Na) in ethanol (200 ml.) in a water-bath at 56° C. for 90 minutes. Water (100 ml.) was then added, the epiphase separated, washed free from alkali with 70 per cent. (v/v) ethanol-water, and finally with water. After drying over  $MgSO_4$ , the solution was chromatographed on a column (56 cm.  $\times$  2.4 cm.) of alumina. The dark orange  $\beta$ -carotene zone was cut, eluted with petroleum ether-methanol and the eluate filtered and concentrated. 42 mg. of crystalline  $\beta$ -carotene was obtained. Some difficulty was experienced in obtaining crystals owing to the presence of the oily substance mentioned, which separated from the mother liquors under the same conditions as the  $\beta$ -carotene. After recrystallization from hexane-methanol, the crystals, dissolved in hexane, showed absorption maxima at 450–1  $m\mu$  and 478  $m\mu$ .

### III. ACKNOWLEDGMENTS

Our thanks are due to Mr. S. A. Clarke, Chief of the Division of Forest Products, and Dr. H. E. Dadswell, Officer-in-Charge, Wood Structure Section, Council for Scientific and Industrial Research, for procuring the log of *Acacia acuminata* and for advice and the provision of facilities; and to Mr. N. L. Lottkowitz for carrying out the microanalysis and microweighings.

### IV. REFERENCES

- (1) TRIKOJUS, V. M., and DRUMMOND, J. C.—*Nature* **139**: 1105 (1937).
- (2) ZECHMEISTER, L., and SANDOVAL, A.—*Arch. Biochem.* **8**: 425 (1945).
- (3) ZECHMEISTER, L., and SANDOVAL, A.—*J. Amer. Chem. Soc.* **68**: 197 (1946).
- (4) KARRER, P., and BENZ, J.—*Helv. Chim. Acta* **31**: 1048 (1948).
- (5) SHANTZ, E. M.—*J. Amer. Chem. Soc.* **68**: 2553 (1946).
- (6) MÜLLER, P. B.—*Helv. Chim. Acta* **26**: 1945 (1943).
- (7) MÜLLER, P. B.—*Ibid.* **27**: 404 (1944).
- (8) MOORE, L. A.—*Industr. Engng. Chem. (Anal. Ed.)* **12**: 726 (1940).
- (9) ZECHMEISTER, L., and POLGÁR, A.—*J. Amer. Chem. Soc.* **65**: 1522 (1943).
- (10) DEVINE, N., HUNTER, R. F., and WILLIAMS, N. E.—*Biochem. J.* **39**: 5 (1945).

### CORRIGENDUM

#### VOLUME 1, NUMBER 2

Pages 233–4, Tables 8, 9A, and 9B, headings to third columns: *For In read log.*



# THE POSITION AND PROBABLE IDENTIFICATION OF THE SOURCE OF GALACTIC RADIO-FREQUENCY RADIATION TAURUS-A

By J. G. BOLTON\* and G. J. STANLEY\*

[*Manuscript received January 18, 1949*]

## *Summary*

This paper contains an account of the discovery of, and subsequent work on, one of the minor discrete sources of galactic radio-frequency noise—Taurus-A. Full details are given of observations and calculations leading to the determination of the source's position and its probable identification with the Crab nebula or supernova of A.D. 1054.

## I. INTRODUCTION

Since Jansky's(1) original discovery of cosmic or galactic radio-frequency noise, considerable information has been accumulated on the intensity distribution of this radiation over the celestial sphere. This distribution roughly follows the optical brightness of the Milky Way, but attempts to account for the intensity and spectrum of the noise in terms of either stellar or interstellar origins have not met with complete success. In 1946 Hey, Parsons, and Phillips(2) found that the noise from the constellation of Cygnus showed short-period variations suggesting that some of the noise from this region was localized in one or several discrete sources. The present writers(3), using an interference technique, showed that these variations originate in an intense discrete source of less than 8' angular width. Later, in a survey over a limited region of the sky, a number of other discrete sources were found(4). They were designated by the number and year of discovery and by a constellation where a rough position could be estimated. Ryle and Smith(5) have since reported the existence of a further intense source in Cassiopeia as well as a number of smaller sources. They also showed that the radiation from the Cygnus source was unpolarized, making it unlikely that the mechanism of the source is similar to that responsible for enhanced solar radio-frequency radiation.

The subject of the present investigation, the source Taurus-A, was first seen on November 6, 1947, during a search for such objects in the Orion region. The technique employed was to observe a fairly extensive region rising above the horizon with an aerial system situated on a high cliff overlooking the sea. The presence of a discrete source was indicated by a lobe pattern formed by interference between the direct ray and the ray reflected from the sea. In Figure 1A, the record obtained on this occasion, the beginning of the lobe pattern of the suspected source is marked by an arrow. The uneven distribution of the general noise of this region passing through the aerial beam produces a varying base level

\* Division of Radiophysics, C.S.I.R.O.

on which the pattern is superimposed. A further series of records obtained in February 1948 showed marked variations on some of the lobes. As the variations always occurred in the same place it was apparent that a second source (8.48) was rising at about the same time as the main source and causing a modulation of the basic lobe pattern. This is illustrated in Figure 1B.

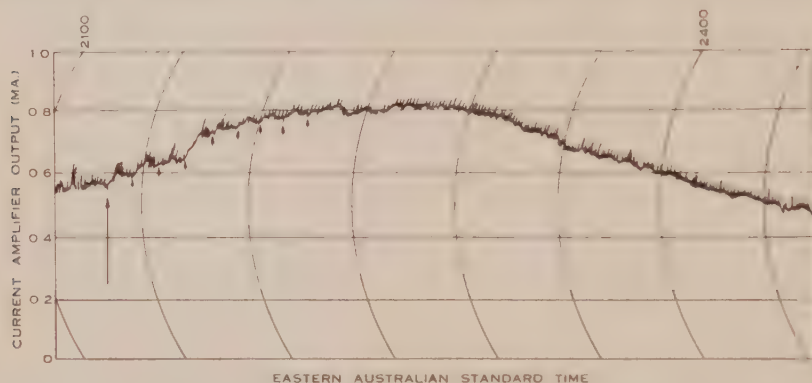


Fig. 1A.—Source Taurus-A as seen from Dover Heights, Sydney, November 6, 1947. Frequency 100 Mc/s. Times, Eastern Australian Standard. Rising point and probable minima are indicated by arrows. Vertical lines on record are due to pick-up from timing mechanism.

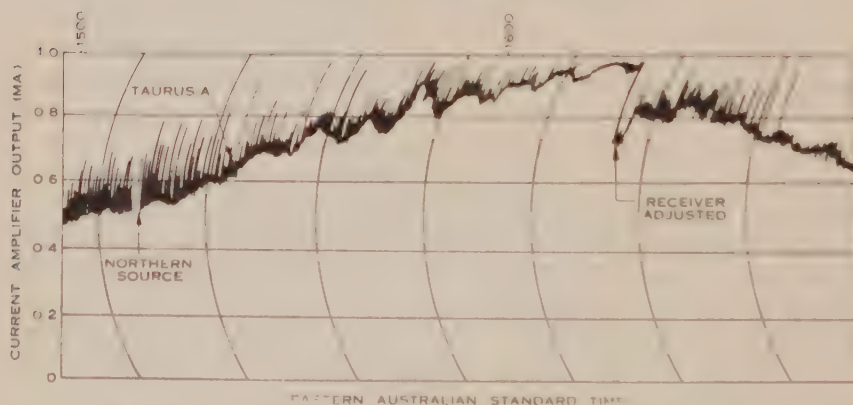


Fig. 1B.—Sources Taurus-A and (8.48) as seen from Dover Heights, Sydney, February 5, 1948. Frequency 100 Mc/s. Times, Eastern Australian Standard. Vertical lines in trace are due to distant atmospherics.

Between June and August 1948 a series of observations was made in New Zealand to determine the exact position of Taurus-A. The sites chosen were Pakiri Hill, near Leigh (on the east coast of the North Auckland peninsula) and a cliff overlooking Piha (on the west coast). Details of sites and equipment used in all observations are given in Table 1. The narrower beamwidth of the aerial system used in New Zealand made it possible to resolve the northern source from Taurus-A. Figure 2 is a record taken at Leigh of the source (8.48) and then Taurus-A rising. It can be seen that after the first four lobes of the Taurus-A

TABLE 1  
SITES AND EQUIPMENT USED IN OBSERVATIONS

Site	Dover Heights, Sydney	Leigh, N.Z.	Piha, N.Z.
Latitude .. .. .	33° 53'	36° 16' 25"	36° 58' 10"
Longitude .. .. .	151° 17'	174° 46' 25"	174° 27' 52"
Height above mean sea-level .. .. .	260 ft.	916 ft.	870 ft.
Frequency .. .. .	100 Mc/s.	100 Mc/s.	100 Mc/s.
Bandwidth .. .. .	1.2 Mc/s.	50 kc/s.	50 kc/s.
Noise factor .. .. .	Nov. 1947—6.0 Feb. 1948—3.7	2.25	2.25
Aerial system .. .. .	1 Yagi	2 Yagis $\lambda$ apart in horizontal plane	
Horizontal beamwidth ..	30°	12°	12°
Vertical beamwidth ..	30°	30°	30°
Post detector amplification factor .. .. .	10	60	20

pattern a modulation effect again occurs, indicating a third discrete source in this small region of the sky. This modulation, in addition to affecting the amplitude of the pattern, displaces the interference minima to a small degree.

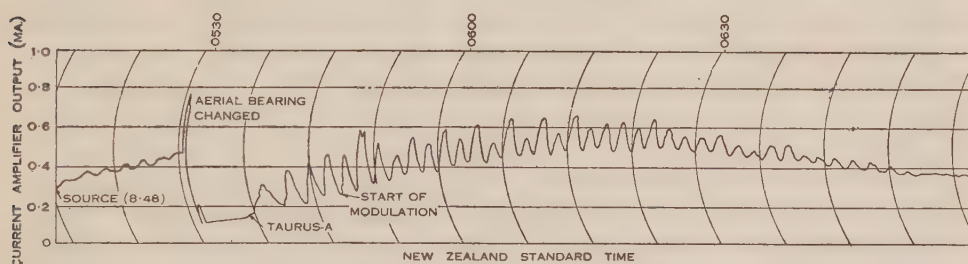


Fig. 2.—Record of sources (8.48) and Taurus-A obtained at Leigh, New Zealand, July 13, 1948. Frequency 100 Mc/s. Times, New Zealand Standard. Note modulation of the Taurus-A interference pattern caused by the third source in this region.

Interference patterns of Taurus-A setting were obtained at Piha but no modulation effect was observed. Owing to solar activity and local interference it was not possible to observe the setting of the two smaller sources.

## II. AN UPPER LIMIT OF ANGULAR WIDTH

From the interference patterns it is possible to place a limit on the size of the source in terms of an "equivalent radiating strip". The full theory of this measurement has been described elsewhere(3, 6). Briefly, however, if  $R$  is the



ratio of the heights of the minima and maxima above the extrapolated cosmic drift level, the width of the equivalent radiating strip  $W$  is given by

$$W = \frac{\lambda}{\pi h} \sqrt{3R},$$

where  $\lambda$  = wavelength,

$h$  = height of aerial above sea-level.

In practice, receiver instability and non-linear drift curves place a limit on the resolving power of this method. With the pattern obtained at Leigh of Taurus-A rising, there are two further difficulties: firstly, the modulation effect occurring after the first ten minutes of the lobe pattern, and secondly, "imperfect" minima due to a sea reflection coefficient of less than unity at apparent angles below  $1^\circ$ . This virtually reduced the number of lobes available for the estimate of  $R$  to one—the third. From a number of records the mean value of  $R$  was less than 0.1, which gives an equivalent radiating strip of less than  $6'$ . From the lobe patterns of the source taken at its setting, a slightly higher upper limit was obtained because recording conditions were unfavourable during daylight. It should be pointed out that measurements made at both setting and rising place an upper limit in the same direction, approximately along the parallel of declination. Further observation will be necessary to ascertain whether this limit has circular symmetry.

### III. DETERMINATION OF POSITION

The method used to determine the position of the source involved estimating its time of rising at one site and its time of setting at the other. From these times was determined the semi-diurnal arc, which in conjunction with the mean latitude of the two observing sites gave an approximate declination. The Right Ascension was derived from the computed time of culmination on the mean longitude of the two sites. Small corrections were then applied to correct the effect of using a mean value for latitude. The procedure adopted to find the necessary rising and setting times is described in subsequent paragraphs, and calculations for the position of the source may be found in the Appendices.

Each minimum on the records of the source rising or setting represents the time at which the source was at an apparent elevation above the horizon. Above  $2^\circ$  the relation between the angle of elevation ( $\alpha$ ), the number of the particular minima ( $n$ ), the wavelength used ( $\lambda$ ), and the height of the aerial system above mean sea-level ( $h$ ) is

$$2h \sin \alpha = n\lambda.$$

Corrections for the effect of earth's curvature (6) are necessary below  $2^\circ$ . The apparent elevations for successive lobes are set out in Tables 2 and 3 together with corrections for atmospheric refraction and derived "true" elevations. Tidal variations which were less than 0.4 per cent. of the heights of the observing sites were neglected and a mean value for the height above sea-level was used.

The times of occurrence of minima on individual days were reduced to a key date for each series by addition or subtraction of the solar-sidereal time correction of  $3' 56''$  per day. These times are also presented in Tables 2 and 3.

Apparent tracks, i.e. plots of elevation against time, were then constructed and are shown in Figure 3 as the right-hand curves of the two series. In the



TABLE 2

TIMES AND ELEVATIONS CORRESPONDING TO MINIMA OF FIVE RECORDS OF TAURUS-A RISING AT LEIGH. KEY DATE JULY 15, 1948. VALUES NOT USED ARE SHOWN IN ITALICS. TIMES ARE EXPRESSED IN MINUTES AND FIFTEENTHS OF A MINUTE, NEW ZEALAND STANDARD TIME

Lobe No.	Times of Minima					Apparent Elevation	Refraction Correction	" True " Elevation (First Approximation)
	July 11	July 12	July 15	July 16	July 18			
	05	05	05	05	05			
0	25/9	23/10	24/7	23/5	25/0	-32'	—	—
1	28/12	28/7	28/2	27/5	28/7	+11'	-51'	-40'
2	31/7	31/7	30/9	30/6	31/3	34'	-44'	-10'
3	33/5	33/8	32/13	32/12	33/8	52'	-39.5'	+12.5'
4	35/9	35/10	35/1	35/4	35/9	1° 11.5'	-35.5'	36'
5	37/13	37/8	36/14	37/5	37/4	1° 31'	-31.5'	59.5'
6	40/3	39/11	39/1	39/11	39/10	1° 50.5'	-28.5'	1° 22'
7	42/2	41/9	41/0	41/12	41/8	2° 10.5'	-26'	1° 44.5'
8	43/8	43/9	42/13	43/13	43/8	2° 29'	-24'	2° 05'
9	45/5	45/8	44/10	45/11	45/7	2° 47.5'	-22'	2° 27.5'
10	47/0	47/5	46/7	47/6	47/3	3° 06.5'	-20.5'	2° 46'
11	49/5	49/3	48/9	49/7	49/4	3° 24'	-19'	3° 05'
12	51/3	51/1	50/3	51/3	51/0	3° 43.5'	-17.5'	3° 26'
13	53/1	53/0	52/9	53/6	53/2	4° 01'	-16.5'	3° 44.5'
14	55/1	54/13	53/13	53/13	54/13	4° 20'	-15.5'	4° 04.5'
15	56/14	56/14	56/2	56/12	56/10	4° 38.5'	-14.5'	4° 24'
17	06 00/13	06 00/10	06 59/2	06 00/11	06 00/7	5° 16'	-13'	5° 03'
20	06/3	06/00	05/10	06/3	06/1	6° 12.5'	-11'	6° 01.5'
22	10/2	10/3	09/8	10/5	09/13	6° 49'	-10'	6° 39'
25	15/10	15/0	15/7	16/1	15/10	7° 44'	-9'	7° 35'
27	19/1	19/9	19/4	19/12	19/6	8° 21.5'	-8.5'	8° 13'
30		25/0	24/13	25/2	25/6	9° 18.5'	-8'	9° 10.5'
32		29/9	28/13	29/7	29/6	9° 50'	-7.5'	9° 42.5'
35		34/10	34/10	34/2	34/5	10° 53.5'	-7'	10° 46.5'
37		39/4	39/0	39/9	39/4	11° 32'	-6.5'	11° 25.5'
40			44/12	43/13	45/9	12° 28.5'	-6'	12° 22.5'

track of the source at its rising, individual points at each elevation are plotted to illustrate the scatter of the results, but only mean points are plotted for the setting track.

As a first approximation in deriving the true elevations, refraction corrections calculated from the Pearcey formula(6) and based on a mean value for the modified refractive index were applied. The resultant corrected tracks are plotted as the second continuous lines in the two series of Figure 3. From the times at which these corrected tracks cut the abscissa for zero elevation an approximate

declination of the source was computed. This approximate value was used to construct theoretical rising and setting tracks for the two sites. These are shown as the dotted curves in Figure 3, plotted in positions convenient for comparison with the "corrected" curves (i.e. slightly displaced). The actual shape of these curves would vary by a negligible amount if the approximate

TABLE 3

TIMES AND ELEVATIONS CORRESPONDING TO MINIMA OF FOUR RECORDS OF TAURUS-A SETTING AT PIHA. KEY DATE AUGUST 4, 1948. VALUES NOT USED ARE SHOWN IN ITALICS. TIMES ARE EXPRESSED IN MINUTES AND FIFTEENTHS OF A MINUTE, NEW ZEALAND STANDARD TIME

Lobe No.	Time of Minima				Apparent Elevation	Refraction Correction	" True " Elevation (First Approximation)
	Aug. 4	Aug. 2	Aug. 3	July 31			
	14	14	14	14			
0	02/3	02/8	00/1	01/11	-30'	—	—
	13	13	13	13			
1	57/0	58/0	55/8	54/11	+11'	-51'	-40'
2	54/0	55/5	51/9	51/11	33'	-44'	-11'
3	51/7	53/0	50/1	49/9	53'	-40'	+13'
4	49/3	50/9	48/1	47/11	1° 14'	-35'	39'
5	46/10	47/8	45/14	45/7	1° 35'	-31'	1° 04'
6	44/14	45/1	43/9	43/11	1° 55'	-28'	1° 27'
7	43/1	42/11	41/13	41/10	2° 16'	-25'	1° 51'
8	41/2	41/0	39/13	39/11	2° 35'	-22.5'	2° 12.5'
9	39/6	38/12	38/1	37/12	2° 55'	-21'	2° 34'
10	37/7	37/2	35/10	35/10	3° 14'	-20'	2° 53'
11	35/12	34/9	33/8	33/14	3° 34'	-19'	3° 15'
12	33/13	32/11	31/8	32/0	3° 54'	-17.5'	3° 36.5'
13	31/14	30/12	29/8		4° 15'	-16'	3° 59'
14	30/0	28/13	27/7		4° 32'	-15'	4° 17'
15		26/9	25/10		4° 51'	-14'	4° 37'
17			21/8		5° 30'	-12.5'	5° 18'
20	16/0		15/7		6° 28'	-10.5'	6° 18'
22			11/9		7° 8'	-10'	6° 58'
25	06/7		08/1		8° 05'	-8'	7° 57'

rising and setting times were in error by as much as one minute, which is unlikely. It will be seen that the theoretical and corrected tracks are closely parallel for relatively high elevations but there is some divergence between them below about 4°, particularly with the rising tracks. It is unlikely that there are serious errors in the corrected tracks at higher angles where the refraction corrections are small. The difference between the theoretical and corrected tracks at high angles was used to adjust the previous values of rising and setting times. The final estimates of these times are shown as the centre of the two blocks on the

time axis of Figure 3. The length of the blocks represents the mean scatter of readings and was used in determining the limits in position of the source.

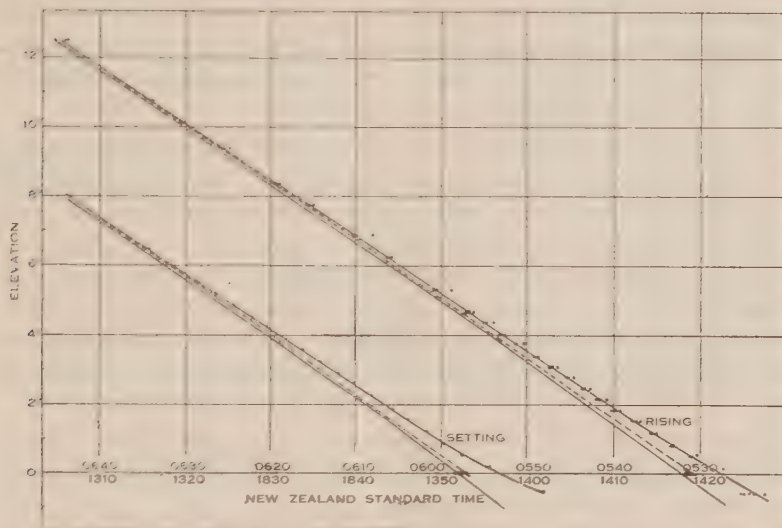


Fig. 3.—Time-elevation curves for the source Taurus-A rising at Leigh (at right) and setting at Piha. From right to left in the two sets the curves are (1) apparent, (2) theoretical, and (3) corrected for refraction following the Pearcey formula. Time runs from right to left for the rising series and left to right for the setting series.

The position of the source Taurus-A as found by this method is :

Right Ascension      5 hr. 31 min. 20 sec.  $\pm 30$  sec.  
Declination             $22^{\circ} 02'$                        $\pm 8'$

#### IV. IDENTIFICATION OF THE SOURCE

The limits in the position of the source enclose NGC 1952, otherwise known as the Crab nebula. According to Baade(7) this nebula is the remains of the supernova of A.D. 1054 observed by Chinese astronomers. The angular dimensions of the nebula are  $4'$  by  $6'$  and the angular rate of expansion  $0.13''$  per year. Doppler shift measurements show an expansion velocity of 1300 km./sec. which with the angular rate gives a distance of 4200 light years.

The colour temperature of the supposed central star is between 7000 and 8000  $^{\circ}\text{K}$ . but a number of factors show that the actual temperature must be much higher :

- (1) The continuous emission spectrum of the photosphere.
- (2) A much greater colour gradient towards the red end of the spectrum than can be explained by interstellar reddening at the known distance of the nebula.
- (3) The total light from the nebula is some seven magnitudes ( $\times 600$ ) greater than that from the central star.

Minkowski(8) considers that most of the radiation from the central star is in the ultraviolet region and the continuous emission from the photosphere is due to bound-free transitions. He estimates a surface temperature of 500,000 °K. and suggests 50,000 °K. for the nebulosity. He explains a low ratio of  $H_{\alpha}$  to [NII] emission by assuming a deficiency of hydrogen in the nebula. This ratio, however, could also be explained by a much higher temperature, which in turn would result in a greater electron density.

The measurements on 100 Mc/s. give an effective temperature of two million degrees, assuming a source size of 5' for Taurus-A. From the present values of temperature and density in the Crab nebula it would be difficult to explain this result in terms of strictly thermal processes. However, it is not unlikely that non-thermal components would arise from differential expansion within the nebula and the general expansion into interstellar matter. In view of this and the close agreement between the positions of the Crab nebula and the source Taurus-A, it is suggested that the Crab nebula is a strong source of radio-frequency radiation.

The regions of the other two supernovae of recent years (Tycho's and Kepler's) have been investigated without result, but there is no visible remnant of the former and very little of the latter.

#### V. ACKNOWLEDGMENT

The work described in this paper was carried out as part of the research programme of the Division of Radiophysics, C.S.I.R.O.

#### VI. REFERENCES

- (1) JANSEY, K. G.—*Proc. Instn. Radio Engrs.* **21**: 1387-98 (1933).
- (2) HEY, J. S., PARSONS, S. J., and PHILLIPS, J. W.—*Nature* **158**: 234 (1946).
- (3) BOLTON, J. G., and STANLEY, G. J.—*Aust. J. Sci. Res. A* **1**: 58 (1948).
- (4) BOLTON, J. G.—*Nature* **162**: 141 (1948).
- (5) RYLE, M., and SMITH, F. G.—*Ibid.* **162**: 462-3 (1948).
- (6) MCCREADY, L. L., PAWSEY, J. L., and PAYNE-SCOTT, RURY.—*Proc. Roy. Soc. A* **190**: 357-75 (1947).
- (7) BAADÉ, W.—*J. Astrophys.* **96**: 188 (1942).
- (8) MINKOWSKI, R.—*Ibid.* **96**: 199 (1942).

#### APPENDIX I

##### *Calculation of Approximate Declination*

	Hr.	Min.	Sec.	
Approximate time of rising at Leigh, July 15, 1948 (N.Z.S.T.) .. ..	05	32	—	(A)
Approximate time of setting at Piha, August 4, 1948 (N.Z.S.T.) .. ..	13	52	—	(B)
Sidereal correction for 20 days .. ..	01	18	38	(C)
Time equivalent of difference in longitudes of two sites .. ..		01	14	(D)
Twice semi-diurnal arc for mean latitude in solar time (B+C-A-D) .. ..	09	37	24	(E)



	Hr.	Min.	Sec.	
Solar-sidereal conversion factor .. ..		1	36	(F)
Semi-diurnal arc for mean latitude				
$[\frac{1}{2}(E+F)]$ .. .. .	04	49	30	
Equivalent hour angle ( $\theta$ ) .. ..		72	23	
Latitude of site at Leigh .. ..	36°	16'	25"	
Latitude of site at Piha .. ..	36°	58'	10"	
Mean latitude ( $l$ ) .. ..	36°	37'	18"	

If  $\delta$  is the declination,

$$\begin{aligned}\tan \delta &= \cos \theta / \tan l \\ &= \cos 72^\circ 23' \\ &\quad \tan 36^\circ 37' \\ \delta &\approx 22^\circ.\end{aligned}$$

## APPENDIX II

### *Right Ascension and Declination using Mean Latitude and Longitude*

The final estimated times of rising and setting of the source are :

	Hr.	Min.	Sec.	
At Leigh, rising time July 15, 1948				
(N.Z.S.T.) .. .. .	05	21	20 $\pm$ 20	
At Piha, setting time August 4, 1948				
(N.Z.S.T.) .. .. .	13	52	20 $\pm$ 40	
Reducing times to July 15				
Rising time, Leigh .. ..	05	31	20 $\pm$ 20	(G)
Setting time, Piha .. ..	15	10	58 $\pm$ 40	(H)
Longitude correction .. ..	00	01	14	(J)
Twice semi-diurnal arc in solar time				
( $H-G-J$ ) .. .. .	09	38	24 $\pm$ 60	(K)
Solar-sidereal time conversion factor	00	01	36	(L)
Semi-diurnal arc for mean latitude				
$[\frac{1}{2}(K+L)]$ .. .. .	04	50	00 $\pm$ 30	
Equivalent hour angle ( $\theta$ ) .. ..	72°	30'	$\pm 7'$	
Mean latitude ( $l$ ) .. ..	36°	37'	18"	
Declination of source ( $\delta$ ) .. ..	22°	02'	$\pm 8'$	

From (G) and (H) the time of culmination on the mean longitude is 10 hr. 21 min. 09 sec.  $\pm$  30 sec. for July 15, 1948 (N.Z.S.T.)

Longitude of site at Leigh .. ..	174°	46'	25"	
Longitude of site at Piha .. ..	174°	27'	52"	
Mean longitude .. ..	174°	37'	09"	
Time equivalent of mean longitude	11 hr.	38 min.	29 sec.	
Time of culmination July 15, 1948 (N.Z.S.T.)	10 hr.	21 min.	09 sec. $\pm$ 30	(M)
Equivalent G.M.T. July 14 .. ..	22	21	09	(N)
Previous transit of Aries (G.M.T.) ..	04	32	11	(P)
Difference in solar time ( $N-P$ ) ..	17	48	58	(Q)
Solar sidereal conversion factor .. ..	00	02	58	(R)
Right Ascension on Greenwich meridian				
when source culminates on mean longitude				
( $Q+R$ ) .. .. .	17	51	56	(S)
Time equivalent of mean longitude ..	11	38	29	(T)
Right Ascension of source ( $S+T$ ) ..	05	30	25 $\pm$ 30	

## APPENDIX III

*Corrections for Using Mean Latitude*

			Hr.	Min.	Sec.	
Semi-diurnal arc used	..	..	04	50	00	
Semi-diurnal arc for Declination $22^{\circ} 02'$						
and latitude of site at Leigh	..	04	50	56	(X)	
Semi-diurnal arc for Declination $22^{\circ} 02'$						
and latitude of site at Piha	..	04	49	06	(Y)	
Semi-sum of (X) and (Y)	..	04	50	01		
Semi-difference of (X) and (Y)	..	00	00	55		

Thus the use of a mean latitude involves a negligible correction for the source's Declination but 55 seconds must be added to the value for the Right Ascension.

*Final Position*

Right Ascension	..	..	..	05 hr. 31 min. 20 sec.	$\pm 30$ sec.
Declination	..	..	..	$22^{\circ} 02'$	$\pm 8'$

# EXPERIMENTAL DESIGNS BALANCED FOR THE ESTIMATION OF RESIDUAL EFFECTS OF TREATMENTS

By E. J. WILLIAMS\*

[Manuscript received December 2, 1948]

## Summary

Where an experiment can be carried out by applying different treatments in succession to the same unit of experimental material, accurate comparisons can be made between the effects of different treatments. To allow for the residual effect of previous treatments on the result obtained for any given treatment, it is desirable to adjust the results for such effects. Methods of constructing balanced designs for the estimation of these residual effects are described in this paper, and are summarized as follows.

*Designs balanced for effect of single preceding treatment:* When  $n$ , the number of treatments, is even, a balanced design is possible with  $n$  replications; when  $n$  is odd,  $2n$  replications are required.

*Designs balanced for the effects of any number of preceding treatments, ignoring the interaction of residual effects:* When  $n$  is a prime or a power of a prime, a balanced design is possible in  $n(n-1)$  replications, which may be set out as a set of  $n-1$  mutually orthogonal Latin squares, with the same first columns. Designs which are not expressible as mutually orthogonal Latin squares are also possible.

*Designs balanced for the effect of the two preceding treatments and their interactions:* A design can be developed from a set of  $n-1$  mutually orthogonal Latin squares obeying certain restrictions.

The method of analysis of designs of this type is set out in detail, together with a numerical example. Direct effects of treatments are shown to be only slightly confounded, the maximum confounding being 4 per cent., when there are three treatments.

These designs have wide applicability wherever successive treatments can be applied to the same unit of experimental material.

## I. INTRODUCTION

Experiments carried out to compare the effects of different treatments on a given material are discussed in this paper. It deals particularly with those cases in which the different treatments can be applied in succession to the same unit of the material. Where this can be done, the treatment effects will be compared as accurately as possible, since the variation between units will not affect the comparisons between the effects of treatments. For this accuracy to be fully realized, however, it is important to make allowance for the effect of previously applied treatments on the response of the unit of material to a given treatment.

The estimation of the effects of previous treatments is of direct interest in experiments, such as crop rotation trials in agriculture, where the aim is to

\* Section of Mathematical Statistics, C.S.I.R.O.

determine the succession of treatments (i.e. crops, fallow, etc.) over a number of years, which will give the best results for the series taken as a whole. This applies, whether the criterion is yield, market value, or the sustained fertility of the soil. On the other hand, such estimation is of indirect interest only, in experiments such as those on the feeding of animals. In such cases, while a group of animals might be tested with a succession of treatments (rations), the primary interest centres in determining the particular ration which gives best results according to some criterion. The effect of previous treatments is of indirect interest, and is estimated only as a correction in determining the effect of the treatment which follows.

The method of carrying out an experiment by applying successive treatments to the same unit of material has wide applicability. The main limiting factors are (i) the length of time taken, compared with simultaneous application of the different treatments to different units of material, and (ii) the suitability of the material to repeated treatments. For example, if the treatment were destructive, the method would not be applicable.

The unit of experimental material to which successive treatments are applied will be called a *site*. The effect of the treatment being applied will be called the *direct effect*, and the effect of previous treatment the *residual effect*, following Cochran, Antony, and Cameron (1). We shall discuss the design of experiments for the estimation of the residual effect, firstly, of the immediately preceding treatment, and secondly, of the two immediately preceding treatments. The analysis of such experiments will be discussed and exemplified only for the former case.

#### (a) *Advantages of a Balanced Design*

We shall call an experimental arrangement a *balanced design* when the direct effect of each treatment is associated equally frequently with the residual effect of each other treatment. Where only the immediately preceding treatment is to be taken into account in assessing residual effects, it is seen that the condition for balance is satisfied if each treatment is preceded equally frequently on the site by each other treatment. Balanced designs have a twofold advantage over unbalanced designs using the same amount of experimental material: firstly, they are more efficient in the sense that they provide more accurate estimates both of direct and residual effects; and secondly, the statistical treatment of the results is relatively simple.

It is therefore of practical interest and importance to determine how balanced designs can be constructed, and what is the minimum number of sites required for balance. Only those designs in which each treatment is applied once to each site will be considered. It would, of course, be possible to construct balanced throughout block or Latin-square designs, with the restriction that they were to be balanced for the estimation of residual effects. Such designs have not been considered, because, firstly, it seems unlikely that balanced designs of this type would be required in practice with so many treatments that all could not be applied at each site; and secondly, the construction and analysis would be more complex than for designs with each treatment at each site.



The designs given in this paper are balanced for the effect of subsequent treatment as well as for previous treatment. Such a possibility is not so academic as it first appears. For example, in carrying out an examination on school children, the result obtained in one test may well be influenced by the anticipation of the subsequent test, to a greater extent, than the actual result of previous test.

## II. BALANCED DESIGN FOR A SINGLE RESIDUAL EFFECT

The conditions for a balanced design are:

- (i) Each treatment shall be preceded by each other treatment equally frequently.
- (ii) Each treatment shall occur equally frequently at each position in order of application to the case so that the treatment effects shall be unaffected by possible effects of order of application.

Each of these conditions implies that the number of cases shall be a multiple of the number of treatments. If there are  $a$  treatments, there are  $a(a-1)$  ordered pairs of treatments. Hence  $a-1$  adjacent pairs occur on each case, resulting in requiring some multiple of  $a$  cases to give all ordered pairs equally frequently. Condition (ii) is satisfied exactly only by a multiple of  $a$  cases.

It will be shown that when  $a$ , the number of treatments, is even, balance can be achieved with a minimum of  $a$  replications, and when  $a$  is odd, with a minimum of  $2a$  replications. In fact, balanced designs can be derived from the cyclic Latin squares of order  $a$  in which the rows represent the cases, the columns the order of application, and the symbols the treatments; two such squares are required when  $a$  is odd. Careful study of Latin squares of order 3 to 6 has revealed no minimum solutions other than those derived from the cyclic squares, and it is surmised that no other solutions exist.

### Case I: $n$ Even

Let each treatment be represented by one of the elements of modulo  $a$  including zero. Condition (i) then can be stated as a requirement that the differences between the values for two adjacent treatments shall take all values from 1 to  $a-1$  equally frequently. If each difference occurs once in the first row of the Latin square, and successive rows are obtained by adding 1 to each element of the preceding row, the values will fulfil the required condition. For each element of any column differs by a constant amount from the corresponding element of the preceding column, the amount being different for each pair of adjacent columns. As each treatment occurs once in each column, it will be preceded on each occasion by a different treatment.

Thus, for example, when  $a=2$ , the differences may be taken as

2   5   3   1   4

giving for the Latin square generated the following arrangement:

0	2	1	4	5	3
1	3	2	5	0	4
2	4	3	0	1	5
3	5	4	1	2	0
4	0	5	2	3	1
5	1	0	3	4	2

It is apparent that each treatment is preceded by each other treatment once and once only.

The problem of determining initial rows of such Latin squares for even values of  $n$  has the following general solution :

$$0 \quad 1 \quad n-1 \quad 2 \quad n-2 \quad 3 \quad n-3 \quad \dots \dots n/2$$

The row is formed from the numbers 0,  $n-1$ , .... in descending order, with which alternate the numbers 1, 2, .... in ascending order. The successive differences are :

$$1 \quad n-2 \quad 3 \quad n-4 \quad \dots \dots \dots ;$$

that is, the odd values in ascending order, alternating with the even values in descending order, and thus include each difference once only. The Latin square developed from this initial row is thus balanced for residual effects of treatments.

Another method of arriving at a balanced arrangement is applicable whenever  $n+1$  is a prime number. In this case, the treatments may be represented by the set of non-zero residue-classes, modulo  $n+1$ . The initial row of the Latin square consists of the residues arranged in numerical order, and successive rows are obtained by multiplying the elements of the initial row in turn by 2, 3, .... to  $n$ . Since  $n+1$  is prime, each treatment will be represented once only in each column. Moreover, the differences of successive elements of any row are all the same as its initial element. Hence, between any treatment  $r$  and the preceding treatment, every difference, except  $r$ , occurs once in the rows which do not begin with  $r$ . Treatment  $r$  is therefore preceded once by each other treatment, and the design is balanced. As an example,  $n=4$  gives

1	2	3	4
2	4	1	3
3	1	4	2
4	3	2	1

At first sight, this method of constructing the balanced square appears to be different from the method previously given. In the first method each difference occurs once in each row, whereas in this method, each difference occurs once between each pair of columns. The two methods can, however, be shown to be equivalent. For, if  $x$  is a residue whose  $n$ th power is the lowest power congruent to 1 modulo  $n+1$ , then all the powers  $x, x^2, \dots, x^n$  are different, and each of the residues 1 to  $n$  can be represented as a power of  $x$ . Multiplying the elements of the first row by the numbers 2 to  $n$  to produce the  $n$  rows of the square, is equivalent to adding a constant amount to the index. Hence, if the treatments are now represented by the indices of  $x$ , the differences between the corresponding elements of adjacent columns will be constant. However, as the design is known to be balanced, it follows that each of the differences from 1 to  $n-1$  must be represented in each row ; where it is to be noted that residues are now to be taken modulo  $n$  and not modulo  $n+1$ . More directly, the last statement follows from the fact that the ratios of successive non-zero residues are all different when the modulus is prime. For example, 2 is a primitive root of the modulus 5, so that

$$\begin{aligned} 1 &= 2^0 \\ 2 &= 2^1 \\ 3 &= 2^3 \\ 4 &= 2^2 \end{aligned}$$

and the above 4-square can be written

0	1	3	2
1	2	0	3
3	0	2	1
2	3	1	0

Another solution, when  $n$  is a power of 2, (say  $2^k$ ), results from giving the successive differences the values

$$1, 2, \dots, n-1.$$

Any cumulative sum of successive differences is of the form

$$\frac{a(a+1)}{2} - \frac{b(b+1)}{2}$$

reduced modulo  $2^k$ , where  $a$  and  $b$  both take values between 0 and  $2^k-1$ . Since the expression equals  $\frac{(a-b)(a+b+1)}{2}$ , and neither  $a-b$  nor  $a+b+1$  is a multiple of  $2^{k+1}$ , the expression is not a multiple of  $2^k$ . Hence, no cumulative sum of successive differences is zero, and these differences provide a solution.

TABLE 1  
INITIAL ROWS FOR BALANCED SQUARES :  $n$  EVEN

$n=4:$	0	1	3	2				
$n=6:$	0	1	5	2	4	3		
	0	2	1	4	5	3		
$n=8:$	0	1	7	2	6	3	5	4
	0	1	3	6	2	7	5	4

Other sets of suitable initial rows have been determined by trial, and are given for  $n=4, 6, 8$ , in Table 1. When one initial row has been found, others can be derived from it by multiplying each element by a residue which is prime to  $n$ , and adding an arbitrary integer. Such rows are equivalent, since the effect of multiplication is the same as permutation of the numbers representing the treatments, and a rearrangement of the rows, and the addition simply permutes the rows. Only a standard representative of each such set of rows is therefore tabulated.

#### Case II: $n$ Odd

When  $n$  is odd, the sum of the residues is a multiple of  $n$ . Hence, if every difference were to occur once in a row, the last element would be the same as the first. It can be seen, therefore, that balanced designs based on a single cyclic Latin square are impossible. It is, however, possible to achieve balance with a pair of such squares, the differences not represented in one square being represented twice in the other.

For example, when  $n=5$ , a possible pair of sets of differences, generating a pair of Latin squares which together provide balance, is

$$1 \ 2 \ 1 \ 3 \quad \quad \quad 2 \ 4 \ 3 \ 4$$

and the corresponding Latin squares are

0	1	3	4	2	0	2	1	4	3
1	2	4	0	3	1	3	2	0	4
2	3	0	1	4	2	4	3	1	0
3	4	1	2	0	3	0	4	2	1
4	0	2	3	1	4	1	0	3	2

It is easy to prove that a pair of Latin squares can be found to give balance for every odd value of  $n$ . The initial row

$$0 \quad 1 \quad n-1 \quad 2 \quad n-2 \quad \dots \quad \frac{n-1}{2} \quad \frac{n+1}{2}$$

gives each odd difference twice, and the reverse arrangement

$$\frac{n+1}{2} \quad \frac{n-1}{2} \quad \dots \quad n-2 \quad 2 \quad n-1 \quad 1 \quad 0$$

gives each even difference twice. The pair of Latin squares generated from this pair of initial rows therefore provides a balanced design.

Another form of general solution is possible when  $n$  is a prime of the form  $4m+3$ . The set of differences

$$1 \quad 2 \quad \dots \quad 2m+1 \quad 2m+1 \quad \dots \quad 2 \quad 1$$

and the opposite set,

$$4m+2 \quad 4m+1 \quad \dots \quad 2m+2 \quad 2m+2 \quad \dots \quad 4m+1 \quad 4m+2$$

together include every difference twice. The rows generated by one set give, on reversal, the rows generated by the other. If the second set of differences is doubled, it becomes

$$4m+1 \quad 4m-1 \quad \dots \quad 3 \quad 1 \quad 1 \quad 3 \quad \dots \quad 4m-1 \quad 4m+1$$

If the middle element of the initial row is taken as zero, the elements to the right are the quadratic residues (modulo  $4m+3$ ), and the elements to the left are their opposites. Hence every residue occurs but once in the row, and the two sets of differences generate a balanced design.

TABLE 2

INITIAL ROWS FOR BALANCED PAIRS OF SQUARES:  $n$  ODD

$n=3:$				0	1	2				0	2	1			
$n=5:$				0	1	3	4	2		0	2	1	4	3	
				0	1	4	2	3		0	4	1	3	2	
$n=7:$				0	1	4	2	6	5	3		0	6	3	5
				0	1	4	3	2	5	6		0	5	2	4
				0	1	6	2	5	3	4		0	6	1	5
				0	1	3	6	2	4	5		0	6	4	1

Pairs of initial rows are given, for  $n=3, 5, 7$  in Table 2. As for even values of  $n$ , other pairs may be derived from these by multiplying each element by a residue prime to  $n$ , and adding an arbitrary integer.



## III. DESIGNS BALANCED FOR PAIRS OF RESIDUAL EFFECTS

It is of both practical and theoretical interest to investigate designs which will give balance for the effects of the two immediately preceding treatments. In deciding on the requirements for these designs, two possibilities arise; we may be prepared to ignore the interaction of one preceding treatment on the other, in assessing their effect on the third, or we may wish to allow for their interaction as well.

In the former case, each treatment must be immediately preceded by each other treatment equally frequently, and must also be preceded by each other treatment by two positions equally frequently. The former condition, as was shown above, requires a multiple of  $n$  replications. As in any set of  $n$  replications, any treatment occupies the first and second positions in order on the site once each, other treatments can occupy the position two before that treatment in only  $n-2$  of the replications. Therefore, in order that each of the  $n-1$  other treatments precede a given treatment by two positions, a multiple of  $n-1$  replications is required. From this it follows that only multiples of  $n(n-1)$  replications are admissible,  $n(n-1)$  being the minimum possible.

In the second case, when interactions are not ignored, it can be seen that each ordered trio of treatments must occur together equally frequently. Now there are  $n(n-1)(n-2)$  such ordered trios, and  $n-2$  occur in each replication. Hence a multiple of  $n(n-1)$  replications is required in this case also.

The latter case, where possible interactions are not ignored, is practically the more useful. It is, however, more restrictive, for designs satisfying the former set of requirements can be found which do not satisfy the latter. The following trio of  $4 \times 4$  Latin squares illustrates this.

0	1	3	2	0	3	2	1	0	2	1	3
1	2	0	3	1	2	3	0	1	0	3	2
2	3	1	0	2	0	1	3	2	3	0	1
3	0	2	1	3	1	0	2	3	1	2	0

It is easily seen that each treatment is preceded by each other treatment, and also preceded by two positions by each other treatment; but, of the  $4 \times 3 \times 2 =$  twenty-four possible ordered trios, six occur twice, twelve occur once, and six not at all.

If designs satisfying the more stringent conditions can be constructed with just  $n(n-1)$  replications, it will be advantageous to use such designs rather than the less restricted ones.

It has been proved(2, 3) that, whenever  $n$  is a prime number or a power of a prime, a completely orthogonalized square of side  $n$  can be constructed. A method of forming a design in  $n(n-1)$  replications, balanced for pairs of residual effects and their interaction will be illustrated, using the properties of the completely orthogonalized square. It is believed that such a design can be constructed whenever a completely orthogonalized square is possible, but we have not been able to prove this. It seems unlikely that such a design in  $n(n-1)$  replications can be constructed other than by such means, such designs being,

therefore, available only when  $n$  is a prime or a power of a prime.\* This is no great disadvantage, however, as, of the numbers from 3 to 9, the likely practicable range, only one is not a power of a prime. It would be no great inconvenience in such cases to add another treatment, or to duplicate one of the existing ones, to bring the number up to a prime power.

It will first be shown that any set of  $n-1$  mutually orthogonal Latin squares gives a design balanced for the effect of any number of preceding treatments, when their interactions are not taken into account, provided that the initial column of each square (or, in fact, any column, the same for each square) has the elements in the same order. This having been proved, it is necessary only to discuss the conditions under which such a set of orthogonal squares is balanced for interactions as well.

Any set of  $n-1$  mutually orthogonal Latin squares may be represented as a set of  $n$  classifications (viz. those corresponding to the rows, and to the symbols of the  $n-1$  squares), which are mutually orthogonal, and orthogonal with columns.

The initial columns of each square may be made identical by suitably numbering the elements of the successive squares, and the rows will be numbered in the same way as the initial columns.

The corresponding symbols for a given column in different squares may now be written down in order, preceded in each case by their row symbol. In this way, the  $n$  columns will give a set of  $n$  squares of side  $n$ . The elements of the first square will be in the same order in each column (since they are the initial columns of the original squares). It will be proved that the remaining  $n-1$  squares are mutually orthogonal Latin squares.

To the original set of Latin squares corresponds a balanced incomplete block design ( $r = n+1$ ,  $v = n^2$ ,  $b = n^2 + n$ ,  $k = n$ ,  $\lambda = 1$ ). There are  $n+1$  complete replications, each comprising  $n$  blocks of  $n$  elements. The first replication may be taken to correspond with the classification of the  $n^2$  elements by columns. The  $n$  elements in the first column will be allocated to different blocks in each of the remaining  $n$  replications. We shall number the blocks in which a given element of the first column occurs, the same in each of these  $n$  replications. This is equivalent to making the initial columns of the original orthogonal squares identical. In each such replication, the  $n$  elements of any column, other than the first, must occur in a differently numbered block. Also any given element of the column must occur in a differently numbered block in each such replication, in order that it may be associated once in the same block with each element of the first column. Hence the square of block numbers, classified according to element of column, and replication, is a Latin square. Furthermore, the elements of any column must be associated once and once only with the elements of each other column. Hence the  $n-1$  Latin squares corresponding to the  $n-1$

\* Since this paper went to press, designs based on sets of cyclic Latin squares which are not orthogonal have been found, including designs for  $n = 6$  and  $n = 10$ . It is intended to describe these in a subsequent note.

columns, are mutually orthogonal. The  $n$  classifications, corresponding to the symbols of the  $n$  squares, are therefore mutually orthogonal.

The formation of this set of orthogonal squares by interchanging columns and squares is illustrated below, for the case  $n=4$ .

		Classification															
Column	..	Row Symbols				Square A				Square B				Square C			
		I	II	III	IV	I	II	III	IV	I	II	III	IV	I	II	III	IV
		1	1	1	1	1	4	2	3	1	3	4	2	1	2	3	4
		2	2	2	2	2	3	1	4	2	4	3	1	2	1	4	3
		3	3	3	3	3	2	4	1	3	1	2	4	3	4	1	2
4	4	4	4	4	1	3	2	4	2	1	3	4	3	2	1		

Corresponding columns of different squares :

Classification	Column I				Column II				Column III				Column IV			
	R	A	B	C	R	A	B	C	R	A	B	C	R	A	B	C
1	1	1	1	1	1	4	3	2	1	2	4	3	1	3	2	4
2	2	2	2	2	2	3	4	1	2	1	3	4	2	4	1	3
3	3	3	3	3	3	2	1	4	3	4	2	1	3	1	4	2
4	4	4	4	4	4	1	2	3	4	3	1	2	4	2	3	1

Since any pair of columns gives rise to an orthogonal pair of classifications in this way, corresponding elements of any pair of columns give every ordered pair of elements (including twins) once. Twins arise only from the row symbols, which are the same for each column. Hence, when these are disregarded, every element is associated with every other element once, in any pair of columns. This establishes the fact that the original set of orthogonal Latin squares gives a balanced design for the effect of any preceding treatments whatever, and consequently for any number of preceding treatments, when their effects are regarded as additive.

There are other possible solutions, however. The illustration given above is an example of this, for it is not a set of orthogonal squares, and cannot be transformed into one by permutation of the rows.

(a) *Sets of  $n-1$  Mutually Orthogonal Latin Squares giving Balanced Designs for Two Preceding Treatments and Their Interaction*

It was shown above that, when interactions are taken into account, a design is required in which every ordered trio occurs equally frequently. The cases when  $n$  is a prime or a power of a prime can both be treated in the same way ; but as the former case is the simpler, and for it a general solution has been found, it will be treated first. For the case when  $n$  is a power of a prime, use is made of the properties of Galois fields.

*Case I:  $n$  an Odd Prime Number*

As before, each treatment is represented by one of the classes of residues, modulo  $n$ , including zero. These classes are the elements of a field, since  $n$  is prime. Suppose that a set of  $n-1$  non-zero differences (not necessarily all distinct) between these elements can be formed, such that (i) no cumulative



sum of successive differences is zero, and (ii) the ratios of consecutive differences take all values except 0,  $n-1$ .

It is noted that, as a result of condition (ii), the first and last differences are equal; for their ratio is the product of all non-zero residues other than  $n-1$ , which is congruent to 1.

If the set of differences be multiplied by 1, 2, . . . . .  $n-1$ , in turn,  $n-1$  sets will be generated, which will comprise between them equally frequently, every non-zero element of the field. That this is so follows from the fact that,  $n$  being prime, any element multiplied in turn by 1, 2, . . . . .  $n-1$ , generates every non-zero element. Moreover, since every such element occurs at each position in the set, in one or other of the sets, it will bear to each difference that precedes it a different ratio (other than  $n-1$ ), so that each will be preceded by each other difference except its opposite, exactly once. In other words, each ordered pair of differences, except those whose sum is zero, will occur once in some set.

Corresponding to each set of differences, a Latin square can be formed, as shown above. The  $n-1$  Latin squares so produced are clearly mutually orthogonal cyclic squares, as their mode of formation is equivalent to that described by Stevens(3). Since every admissible ordered pair of differences is adjoined to each element by this means, in one or other of the squares, every ordered trio of elements is thereby produced. The set of Latin squares therefore gives the required balanced design.

The method of formation is made clear by means of an example, for  $n=5$ .

The set of differences

1 3 3 1

satisfies the conditions (i) and (ii) above; the successive ratios are

3 1 2

and the ratio 4 is inadmissible.

Three other sets of differences may be produced by multiplying the given set by 2, 3, and 4, successively; the four sets are

1 3 3 1  
2 1 1 2  
3 4 4 3  
4 2 2 4

From these four sets, the following four orthogonal Latin squares are derived. Every ordered trio is seen to be represented once.

0 1 4 2 3  
1 2 0 3 4  
2 3 1 4 0  
3 4 2 0 1  
4 0 3 1 2

0 2 3 4 1  
1 3 4 0 2  
2 4 0 1 3  
3 0 1 2 4  
4 1 2 3 0

0 3 2 1 4  
1 4 3 2 0  
2 0 4 3 1  
3 1 0 4 2  
4 2 1 0 3

0 4 1 3 2  
1 0 2 4 3  
2 1 3 0 4  
3 2 4 1 0  
4 3 0 2 1



For any odd prime number, a set of differences satisfying conditions (i) and (ii) is

$$1 \quad n-2 \quad 3 \quad n-4 \quad \dots \quad \frac{n+1}{2} \quad \frac{n+1}{2} \quad \dots \quad n-2 \quad 1$$

and the corresponding initial row of the first Latin square of the set is

$$0 \quad 1 \quad n-1 \quad 2 \quad n-2 \quad \dots \quad \frac{n-1}{2} \quad \frac{n+1}{2}$$

The  $n(n-1)$  rows obtained by multiplying the elements of the initial row by 1, 2, 3, . . . . .  $n-1$ , and in each case adding to the resulting elements 0, 1, 2, . . . . .  $n-1$  give the required design balanced for pairs of residual effects. This proves the existence of such a design when  $n$  is any odd prime.

This general solution is the same as that found above for a single Latin square giving balance for the effect of one preceding treatment only, when  $n$  is even, or for the pair of Latin squares required when  $n$  is odd.

TABLE 3

DESIGNS BALANCED FOR PAIRS OF RESIDUAL EFFECTS.  
INITIAL ROWS FOR GENERATING SETS OF ORTHOGONAL  
SQUARES.  $n$  A PRIME

$n=3:$	0	1	2				
$n=5:$	0	1	4	2	3		
$n=7:$	0	1	2	6	5	3	4
	0	1	6	5	4	2	3
	0	1	6	2	5	3	4
	0	1	3	6	2	4	5
	0	1	4	3	2	5	6

Other initial rows also give balanced designs in the same way; these are tabulated in Table 3, for  $n=3, 5, 7$ .

#### Case II: $n$ a Power of a Prime

When  $n=p^k$ , where  $p$  is prime and  $k$  an integer, each treatment may be represented by one of the elements of a Galois field. Galois fields exist only for  $n$  a prime or a power of a prime. A full treatment of the properties of these fields will be found in Carmichael (4, pp. 242-88).

Every field has a primitive element  $x$ , with the property that the equation  $x^m=1$  is satisfied only by  $m$  a multiple of  $n-1$ . Consequently, the elements  $x, x^2, x^3, \dots, x^{n-1}$  ( $=1$ ) are all distinct, and therefore include every non-zero element of the field.

It can be shown that such a primitive element satisfies an equation  $F(x)=0$ , where  $F(x)$  is an irreducible polynomial of degree  $k$ , whose coefficients are members of  $GF(p)$ .  $F(x)$  is called a minimum function. It follows that every element of the field may be expressed as a polynomial in  $x$  of degree  $k-1$ , namely

$$c_0 + c_1x + \dots + c_{k-1}x^{k-1},$$

where the coefficients  $c_0, c_1, \dots, c_{k-1}$  are members of  $GF(p)$ . Addition and multiplication are defined as ordinary addition and multiplication, followed by simultaneous reduction modulo  $p$  and modulo  $F(x)$ .

We note that, when  $n$  is the power of an odd prime,

$$x^{n-1} - 1 = 0$$

but

$$x^{\frac{n-1}{2}} - 1 \neq 0;$$

therefore

$$x^{\frac{n-1}{2}} + 1 = 0.$$

When  $n$  is a power of 2,

$$1 + 1 = 0.$$

We are now in a position to adapt the procedure used when  $n$  is prime, to the derivation of a balanced design when  $n$  is a power of a prime. We require to find a set of  $n-1$  non-zero differences among the elements of the field, such that (i) no cumulative sum of successive differences is 0, and (ii) the ratios of successive differences take all values except 0,  $x^{\frac{n-1}{2}}$  (when  $n$  is odd) and all values except 0, 1 (when  $n$  is a power of 2).

As for the case of  $n$  a prime, the first and last differences must be equal.

If the set of differences be multiplied by 1,  $x, x^2, \dots, x^{n-2}$ , in turn,  $n-1$  sets will be generated. Corresponding to each such set of differences, a Latin square can be formed. The argument given above for the case when  $n$  is prime can be extended to show that the Latin squares so constructed are mutually orthogonal, and that every ordered trio of elements is represented once. This set of squares therefore gives the required balanced design.

Although it is believed that a solution exists for every  $n$  which is a power of a prime, no general solution has been found. Solutions for the particular cases likely to be required in practice, namely,  $n = 4, 8, 9$ , are, however, presented in Table 4. The initial row is given, also the minimum function and the power cycle for the primitive root  $x$ . The initial row is standardized by choosing 0 and 1 as the first two elements.

From any given solution, other solutions may be derived by raising each element to a power whose index is a power of  $p$ . The addition and multiplication relationships among the new elements are exactly the same as those among the old. For the  $p$ th power of a sum is congruent (modulo  $p$ ) to the sum of the  $p$ th powers; while the  $p$ th power of a product is obviously identical with the product of  $p$ th powers. There are just  $k$  different solutions so derivable; for if  $x$  is any element,

$$x^{p^k} = x^n = x.$$

This fact enables initial rows to be further standardized. For example, when  $n=9$ , the following two initial rows are equivalent:

$$\begin{array}{cccccccc} 0 & 1 & x & x^6 & x^3 & x^3 & x^4 & x^7 & x^6 \\ 0 & 1 & x^3 & x^7 & x^6 & x & x^4 & x^5 & x^3 \end{array} \quad .$$

In Table 4, the solution is given for which the third element is the lowest possible power of  $x$ .

As an example of the generation of a balanced design from the initial row given in Table 4, the complete design is given for  $n=4$ . In setting out the design in this way, it is convenient to represent the zero element by 0, the unit element by 1, and the successive powers of  $x$  by 2, 3, . . . . .  $n-1$ .

0	1	2	3	0	2	3	1	0	3	1	2
1	0	3	2	1	3	2	0	1	2	0	3
2	3	0	1	2	0	1	3	2	1	3	0
3	2	1	0	3	1	0	2	3	0	2	1

TABLE 4

DESIGNS BALANCED FOR THE EFFECTS OF PAIRS OF RESIDUAL EFFECTS. INITIAL ROWS FOR GENERATING SETS OF ORTHOGONAL SQUARES WHEN  $n$  IS A POWER OF A PRIME

$n$	Initial Row	Minimum Function	Power Cycles
4	0 1 $x$ $x^2$	$x^2+x+1$	$x=x, x^2=x+1, x^3=1$
8	0 1 $x$ $x^6$ $x^3$ $x^2$ $x^5$ $x^4$ 0 1 $x$ $x^3$ $x^5$ $x^4$ $x^2$ $x^6$ 0 1 $x^3$ $x$ $x^4$ $x^5$ $x^2$ $x^6$ 0 1 $x^3$ $x^4$ $x$ $x^5$ $x^6$ $x^2$	$x^3+x+1$	$x=x, x^2=x^2, x^3=x+1,$ $x^4=x^2+x, x^5=x^2+x+1,$ $x^6=x^2+1, x^7=1$
9	0 1 $x$ $x^5$ $x^2$ $x^3$ $x^4$ $x^7$ $x^6$ 0 1 $x$ $x^5$ $x^7$ $x^4$ $x^6$ $x^2$ $x^3$	$x^2+x+2$	$x=x, x^2=2x+1, x^3=2x+2$ $x^4=2, x^5=2x, x^6=x+2$ $x^7=x+1, x^8=1$

#### IV. ANALYSIS OF RESULTS, TAKING INTO ACCOUNT RESIDUAL EFFECT OF PRECEDING TREATMENT

There are four factors to be taken into account in the analysis; sites or replications, position in order of application to the site, and direct and residual effects of treatments. It is noted first of all that, since each treatment occurs equally frequently at each position, positions are orthogonal with each of the other effects. The determination of the effect of the positions therefore is made direct from their means without adjustment. Sites and direct effects are also orthogonal, but neither is orthogonal with residual effects. The analysis ignoring residual effects is straightforward, and follows the lines of the analysis of a Latin square or set of Latin squares. When residual effects are taken into account, the analysis is more complex, but as the design is balanced, the computations are not difficult.

It is useful to note that sites bear the same relationship to residual effects as do direct effects. For consider all the sites in which treatment  $i$  occupies the final position. On these sites, the residual effect of treatment  $i$  alone is absent. Comparisons among such sites, in which the same treatment comes last, are therefore orthogonal with residual effects of treatments. On the other hand, comparisons between the groups of sites with different final treatments are not

orthogonal with residual effects. Since treatment  $i$  is never preceded by treatment  $i$ , the total for any treatment is affected by residual effects in exactly the same way as is the total for any set of replications for which that treatment comes last.

(a) *Notation for the Description of the Analysis*

Totals	{	$n$	=number of treatments,
		$mn$	=number of sites or replications,
		$S_{ik}$	=total of results for the $k$ th site of those for which treatment $i$ is the final one,
		$F_i$	=total of results for all sites for which treatment $i$ is the final one (designated group $i$ ),
		$T_i$	=total of results for treatment $i$ ,
		$R_i$	=total of results for treatments preceded by treatment $i$ ,
		$P_j$	=total of results for position $j$ in order of application to the site,
		$G$	=grand total of all results.
Estimates (measured as deviations from zero)	{	$f_i$	=average effect for all sites of group $i$ ,
		$t_i$	=direct effect of treatment $i$ ,
		$r_i$	=residual effect of treatment $i$ ,
		$r'_i$	=residual effect of treatment $i$ (direct effects being ignored),
		$p_j$	=effect of position $j$ in order of application.
		$i$	=1, 2, . . . . . $n$
		$j$	=1, 2, . . . . . $n$
		$k$	=1, 2, . . . . . $m$ .

(b) *Derivation of the Estimates, and Tests of Significance*

The normal equations are

$$m(nf_i - r_i) = F_i - \frac{G}{n},$$

$$m(nt_i - r_i) = T_i - \frac{G}{n},$$

$$m((n-1)r_i - t_i - f_i) = R_i + \frac{P_1}{n} - \frac{G}{n},$$

$$mnp_j = P_j - \frac{G}{n}.$$

The solutions of the normal equations are

$$f_i = \frac{(n^2 - n - 1)F_i + nR_i + T_i + P_1 - nG}{mn(n^2 - n - 2)}$$

$$t_i = \frac{(n^2 - n - 1)T_i + nR_i + F_i + P_1 - nG}{mn(n^2 - n - 2)}$$

$$r_i = \frac{n^2R_i + nF_i + nT_i + nP_1 - (n+2)G}{mn(n^2 - n - 2)}$$

$$r'_i = \frac{n^2R_i + nF_i + nP_1 - (n+1)G}{mn(n^2 - n - 1)}$$



$$p_j = \frac{nP_j - G}{mn^2}.$$

The estimates given here take into account only the information given by comparisons within sites (actually, within groups of sites with the same final treatment). If the information provided by comparisons between groups of sites be taken into account, slightly more accurate estimates result.

It can be shown that residual effects are confounded  $1/(n^2-n)$  with groups of sites and with treatments. Since only that part confounded with sites is potentially recoverable, the fraction of recoverable information confounded with sites is  $1/(n^2-n-1)$ .

In the same way it is found that direct effects of treatments are confounded  $1/(n^2-n)$  with residual effects when sites are not taken into account. The fraction thus confounded is of course not recoverable, so that the fraction potentially recoverable is  $(n^2-n-1)/(n^2-n)$ . When sites are taken into account, an additional fraction  $1/((n^2-n)(n^2-n-1))$  is confounded. Hence, the fraction of potentially recoverable information on direct effects, confounded with sites, is  $1/(n^2-n-1)^2$ . The efficiency factors for estimates from comparisons within sites are therefore as follows:

$$\begin{aligned} \text{Direct effects} \quad & \frac{(n+1)n(n-1)(n-2)}{(n^2-n-1)^2} = 1 - \frac{1}{(n^2-n-1)^2} \\ \text{Residual effects} \quad & \frac{(n+1)(n-2)}{(n^2-n-1)} = 1 - \frac{1}{(n^2-n-1)} \end{aligned}$$

TABLE 5  
EFFICIENCY FACTOR OF ESTIMATES OF DIRECT AND RESIDUAL  
EFFECTS

<i>n</i>	<i>E</i> (%)	
	Direct	Residual
3	96.0	80.0
4	99.2	90.9
5	99.7	94.7
6	99.9	96.6

These factors are presented, for small values of *n*, in Table 5. Unless *n* is small, the gain due to recovery of information from comparisons among sites will be slight, and will moreover be offset to some extent by the loss of information due to inaccuracies in weighting. However, recovery of information on residual effects may sometimes be worth while, when *n* is 3 or 4.

In many cases, where the estimation of the residual effect is not of direct interest, and it is required only as an adjustment to the average effect of a

treatment, estimates of the residual effect with improved accuracy are irrelevant. The recovery of information is, however, discussed below, to meet the cases in which the residual effects are of direct interest.

Ordinarily, two analyses of variance will be required for testing the significance of treatments and of the adjustment for residual effects. The sum of squares for direct effects, corrected for residual effects, and the sum of squares for residual effects, corrected for direct effects, will both need to be calculated. The two analyses are set out in Table 6.

TABLE 6

	Degrees of Freedom	Sum of Squares	
		Direct Effects Adjusted	Residual Effects Adjusted
Sites—Between groups	$n-1$	$\frac{\sum F_i^2}{mn} - \frac{G^2}{mn^2}$	
Within groups ..	$n(m-1)$	$\frac{\sum \sum S_{ik}^2}{n} - \frac{\sum F_i^2}{mn}$	
Positions .. ..	$n-1$	$\frac{\sum P_j^2}{mn} - \frac{G^2}{mn^2}$	
Direct effects .. ..	$n-1$	$\frac{\sum (mn(n^2-n-2)t_i)^2}{mn(n^2-n-1)(n^2-n-2)}$	$\frac{\sum T_i^2}{mn} - \frac{G^2}{mn^2}$
Residual effects ..	$n-1$	$\frac{\sum (m(n^2-n-1)r'_i)^2}{mn(n^2-n-1)}$	$\frac{\sum (m(n^2-n-2)r_i)^2}{mn(n^2-n-2)}$
Error .. .. .	$(mn-3)(n-1)$	by subtraction	
Total .. .. .	$mn^2-1$	dev <sup>2</sup> y	

(c) *Standard Errors of Estimates of Treatment Effects*

The estimated variance of the difference between two estimates of direct effects, when residual effects are ignored, is

$$\frac{2E}{mn},$$

where  $E$  is the error mean square from the analysis of variance. When the adjustment for residual effects is made, the variance of a difference is increased to

$$\frac{2E}{mn} \frac{n^2-n-1}{n^2-n-2}.$$

The variance of the difference between the estimates for two residual effects is

$$\frac{2nE}{m(n^2-n-2)}.$$

(d) *Recovery of Information between Groups of Sites*

Let the error variance be  $1/W$ , and the variance between sites of the same group, on a single-result basis, be  $1/W'$ .

If  $E$  is the error mean square and  $S$  the mean square between sites within groups, their expected values are respectively  $1/W$  and  $1/W'$ . However, if  $E$  exceeds  $S$ , we assume  $W$  and  $W'$  to be equal, and take  $1/E$  as the estimate of each.

The estimate of the residual effect of treatment, after including information from groups of sites, is

$$\begin{aligned}\hat{r}_i &= \frac{r_i(n^2-n-2)S - \frac{nF_i - G}{mn}E}{(n^2-n-2)S + E} \\ &= \frac{(n^2R_i + nF_i + nT_i + nP_1 - (n+2)G)S - (nF_i - G)E}{mn((n^2-n-2)S + E)}.\end{aligned}$$

The estimated variance of the difference between two such estimates is

$$\frac{2nSE}{m((n^2-n-2)S + E)} = \frac{2nE}{m(n^2-n-2)} \left\{ 1 - \frac{E}{(n^2-n-2)S + E} \right\}.$$

The expression for the variance of differences may be compared with that obtained above, without recovery of information. The expression in brackets represents the reciprocal of the relative efficiency with recovery of information.

TABLE 7  
ORDER OF APPLICATION OF CONCENTRATIONS IN  
SUCCESSIVE REPLICATIONS

Replication						
I	3	6	2	5	4	1
II	5	3	4	6	1	2
III	1	4	5	2	6	3
IV	2	1	6	4	3	5
V	6	5	1	3	2	4
VI	4	2	3	1	5	6

*Numerical Example*

The example given below is taken from experimental pulp evaluation technique. Samples of pulp suspension at varying concentrations are beaten in a Lampén mill, to determine the effect of concentration on the properties of the resulting sheets. Observations of the condition of the mill after each beating indicate that certain concentrations of pulp have an effect on the mill which may

affect the next beating. In order to allow for this effect if it exists, a balanced design is adopted. There are six concentrations used, in all, and six replications are carried out. The layout of the experiment is shown in Table 7 and the results obtained for the burst factor of the resulting sheets in Table 8.

TABLE 8  
BURST FACTOR RESULTS FOR THE EXPERIMENT OF TABLE 7

Replication							Total
I	56.7	53.8	54.4	54.4	53.9	54.5	332.7
II	58.5	60.2	61.3	54.4	59.1	59.8	353.3
III	55.7	60.7	56.7	59.9	56.6	59.6	349.2
IV	57.3	57.7	55.2	58.1	60.2	60.2	348.7
V	53.7	57.1	59.2	58.9	58.9	59.6	347.4
VI	58.1	55.7	58.9	56.6	59.6	57.5	346.4
Total	340.0	345.2	345.7	342.3	353.3	351.2	2077.7

TABLE 9  
TOTALS AND SUMS OF SQUARES REQUIRED FOR ANALYSIS

			$A_i$	$B_i$	$C_i$	$D_i$
In terms of estimates			$36r'_i$	$174r'_i$	$168r_i$	$168r_i$
In terms of totals $i$	$R_i$	$F_i$	$6T_i - G$	$36R_i + 6F_i + 6P_1 - 7G$	$B_i + A_i$	$\frac{1}{2}(C_i + 28A_i)$
1	294.2	332.7	-20.9	83.5	62.6	-87.1
2	287.2	353.3	-1.7	-44.9	-46.6	-15.7
3	290.8	349.2	49.3	60.1	109.4	248.3
4	281.5	347.4	62.5	-285.5	-223.0	254.5
5	295.7	348.7	1.3	233.5	234.8	45.2
6	288.3	346.4	-90.5	-46.7	-137.2	-445.2
Add $P_1$	1737.7	2077.7	0.0	0.0	0.0	0.0
	340.0					
	2077.7					
Divisor for sum of squares	..	..	216	6264	6048	4872
Sum of squares	..	..	69.298	24.076	23.436	68.658



For numerical work it is convenient to use the quantities given below in calculating the estimates and the sums of squares in the analysis of variance. Corresponding to any treatment, all but the first of these quantities may be

TABLE 10  
ANALYSIS OF VARIANCE

	Degrees of Freedom	Sum of Squares		Mean Square (Adjusted)
		Direct Effects Adjusted	Residual Effects Adjusted	
Replications ..	5	41.558		
Positions in order of beating ..	5	21.711		4.342
Direct effects ..	5	68.658	69.298	13.73**
Residual effects	5	24.076	23.436	4.687**
Error .. ..	15	9.059		0.6039
Total ..	35	165.062		

\*\* Significant at the 1% level.

calculated successively on the machine without clearing out the previously obtained total. In the final column is shown the divisor for the sum of squares of the quantities in the previous column, to give the appropriate sum of squares for the analysis of variance.

TABLE 11  
ADJUSTED TREATMENT MEANS AND ESTIMATES OF RESIDUAL EFFECTS

Treatment	Adjusted Mean	Residual Effect
1	57.20	0.37
2	57.62	-0.28
3	59.19	0.65
4	59.23	-1.33
5	57.98	1.40
6	55.06	-0.82
Standard error of difference	0.46	0.51

	Divisor
$A_i = mn^2 t'_i = nT_i - G$	$mn^3$
$B_i = mn(n^2 - n - 1)r'_i = n^2 R_i + nF_i + nP_1 - (n+1)G$	$mn^3(n^2 - n - 1)$
$C_i = mn(n^2 - n - 2)r_i = n^2 R_i + nF_i + nT_i + nP_1 - (n+2)G$ $= B_i + A_i$	$mn^3(n^2 - n - 2)$
$D_i = mn(n^2 - n - 2)t_i = (n^2 - n - 1)T_i + nR_i + F_i + P_1 - nG$ $= \frac{1}{n}(C_i + (n^2 - n - 2)A_i).$	$mn(n^2 - n - 1)(n^2 - n - 2)$

The detailed analysis of the results of the experiment is set out in Tables 9, 10, and 11.

#### V. ACKNOWLEDGMENT -

The work described in this paper was carried out as part of the research programme of the Section of Mathematical Statistics, C.S.I.R.O.

#### VI. REFERENCES

- (1) COCHRAN, W. G., AUTREY, K. M., and CANNON, C. Y.—A double change-over design for dairy cattle feeding experiments. *J. Dairy Sci.* **24** : 937 (1941).
- (2) BOSE, R. C.—On the application of the properties of Galois fields to the problem of construction of Hyper-Graeco-Latin squares. *Sankhyā* **3** : 323 (1938).
- (3) STEVENS, W. L.—The completely orthogonalized Latin square. *Ann. Eugen.* **9** : 82 (1939).
- (4) CARMICHAEL, R. D.—“Introduction to the Theory of Groups of Finite Order.” (Ginn and Co. : Boston and London, 1937.)

# THE WAVE EQUATIONS FOR ELECTROMAGNETIC RADIATION IN AN IONIZED MEDIUM IN A MAGNETIC FIELD

By K. C. WESTFOLD\*

[Manuscript received January 5, 1949]

## Summary

The inadequacy of the present practice of applying the formulae of magneto-ionic theory to ray trajectories satisfying Snell's law, to represent the conditions of propagation of radio waves in the ionosphere and the solar atmosphere, is explained.

To express the differential equations of the electromagnetic field as simply as possible the axes of coordinates are chosen to coincide with the three principal directions of the medium. Wave equations for the components of the electric field intensity and the electric Hertz vector are investigated.

In applying the theory to the propagation of plane waves in a uniform medium, a new fundamental form of the equation for the complex refractive index of the Appleton-Hartree theory is obtained. The three types of electron plasma oscillations possible in such a medium are also discussed.

It is suggested that solar "bursts" should decay according to a damping constant of the order of the collision frequency where they originate in the solar atmosphere.

## I. INTRODUCTION

Many authors have discussed the extension, to more general conditions of propagation, of Appleton's classical magneto-ionic theory, which is adequate for the description of the propagation of plane waves in a *uniformly* ionized medium in a *uniform* magnetic field.

The first wave theory developments by Hartree(1) and later by Booker(2) cannot be regarded strictly as extensions of magneto-ionic theory, because Hartree and Booker assumed all the electromagnetic properties of the medium to be stratified with respect to the electromagnetic radiation which can be propagated within it. The refractive index is then a function of one coordinate only. That this is not generally true can be seen by consideration of the behaviour of a plane wave which enters an ionized medium having a slowly varying stratified distribution of electron density in a uniform magnetic field. This is the standard model for an ionospheric region.

In general, since there are two alternative modes of propagation in an ionized medium in a magnetic field (in future we shall describe such a medium as "*magneto-ionic*"), we may assume that the incident wave will separate into two corresponding waves, although we have no criterion as to how the energy is distributed between them.

Ray methods consist essentially of the application of Fermat's principle to the trajectory of a limited portion of a wave-front. The trajectory and the

\* Division of Radiophysics, C.S.I.R.O.

absorption and polarization changes suffered by the ray are determined by the refractive index structure of the medium. If surfaces of constant refractive index are determinate, Fermat's principle reduces to the simple form of Snell's law. It is the usual practice to regard the surfaces of constant electron density in a magneto-ionic medium as equivalent to surfaces of constant refractive index.

However, according to the Appleton-Hartree formula, the refractive index is in general a function of the angle the direction of propagation makes with the direction of the magnetic field as well as of the electron density, the magnetic field intensity, and the frequency. The differential deviation of a ray at any point of its trajectory depends on all these parameters, so that the difference in refractive index at neighbouring points on a ray in the medium under consideration depends on the deviation as well as on the new value of the electron density. The refractive index structure of the medium is therefore an intrinsic property of the particular ray under consideration. It follows that surfaces of constant refractive index, independent of ray directions, cannot be realized in a magneto-ionic medium. This is the fundamental difficulty which besets all attempts at a ray treatment of the propagation of radio waves through the ionosphere or the solar atmosphere.

Since the simplification which stratification would offer is denied us, the only remaining method of simplification is to choose the coordinate framework of the theory in the most suitable way. When plane waves are being considered there are two choices: the first, made by Appleton(3), is such that one of three rectangular axes is parallel to the direction of propagation, the second, made by Nichols and Schelleng(4), is such that one of the three rectangular axes is parallel to the direction of the imposed magnetic field. Most subsequent writers have followed Appleton's choice, though it does not seem well suited to waves whose direction of propagation changes as they travel in the medium.

The second choice is not subject to this criticism. It was made by Baker and Green(5), and advocated by Huxley(6), Saha and Banerjee(7), and Banerjee(8). Saha and Banerjee demonstrated the fundamental nature of this choice by showing that the dielectric tensor, which specifies the properties of the medium, had the real characteristic principal axis of its associated tensor ellipsoid parallel to the direction of the imposed magnetic field. They proceeded to develop the electromagnetic field equations in these coordinates, but made the error of assuming that the divergence of the electric field was zero instead of the divergence of the electric displacement. In a more recent paper by Saha, Banerjee, and Guha(9) the correct field equations are stated at the outset but the wave equations they obtain for the field components are the same as in the earlier paper. A striking result of this error is that when their equations are applied to plane waves propagated in a direction normal to the imposed magnetic field, in a uniform medium, the "extraordinary" wave of classical magneto-ionic theory cannot be deduced.

The chief characteristic of a magneto-ionic medium is its anisotropy. Following the usual treatment of crystal optics, in this paper we investigate the three principal directions of such a medium and develop component wave



equations of the electromagnetic field using a coordinate system of axes parallel to these directions.

For the simpler applications of the theory, it is sufficient to confine attention to the electric vector  $\mathbf{E}$ , but if the source of the electromagnetic field is to be specified or if the initial or boundary conditions involve the magnetic vector  $\mathbf{H}$  as well as  $\mathbf{E}$ , the representation of the electromagnetic field in terms of the electric Hertz vector  $\mathbf{Z}$  is most appropriate. We therefore investigate both developments of the wave equations.

Of the simpler cases we then investigate the characteristics of plane waves and plasma oscillations which can exist in a magneto-ionic medium.

## II. THE FIELD EQUATIONS

In terms of rationalized M.K.S. units, the equations of the electromagnetic field intensities  $\mathbf{E}$  and  $\mathbf{H}$  are

$$\left. \begin{aligned} \text{curl } \mathbf{E} + \mu_v \frac{\partial \mathbf{H}}{\partial t} &= 0, \\ \text{curl } \mathbf{H} - \frac{\partial \mathbf{D}}{\partial t} &= 0. \end{aligned} \right\} \dots\dots\dots (2.1)$$

The electric displacement vector  $\mathbf{D}$  is related to  $\mathbf{E}$  by the equation

$$\mathbf{D} = \epsilon_v \mathbf{E} + \mathbf{P} \dots\dots\dots (2.2)$$

where  $\mathbf{P}$  is the electric polarization vector whose relation to  $\mathbf{E}$  is given by the Lorentz force-equation

$$\frac{\partial^2 \mathbf{P}}{\partial t^2} + \nu \frac{\partial \mathbf{P}}{\partial t} = \frac{Ne^2}{m} \mathbf{E} + \frac{e\mu_0}{m} \frac{\partial \mathbf{P}}{\partial t} \times \mathbf{H}_0 \dots\dots\dots (2.3)$$

In (2.2) and (2.3),  $\epsilon_v$  and  $\mu_v$  are respectively the permittivity and permeability of free space,\*  $N$  is the electron density,  $e$  and  $m$  the charge and mass of an electron,  $\nu$  the mean frequency of elastic collisions made by an electron with the other particles of the medium, and  $\mathbf{H}_0$  the imposed magnetic field. In (2.3) we have neglected to follow the motion of the electrons and assumed that the magnetic vector  $\mathbf{H}$  is small compared with  $\mathbf{H}_0$ . In applications to the ionosphere and the solar atmosphere these assumptions are justified and have the effect of making the relation linear. Bailey(10) has recently developed an "electromagneto-ionic" theory where in addition to  $\mathbf{H}_0$ , a constant electric field  $\mathbf{E}_0$  is imposed on the medium. The effect of following the motion of the electrons then becomes important because of the drift velocities they attain in the field  $\mathbf{E}_0$ . We have also neglected the contributions made to  $\mathbf{P}$  by the far heavier positive ions and, following Darwin(11), have taken the Lorentz term to be zero.

Taking the divergence of both equations (2.1), we get the remaining Maxwell relations

$$\left. \begin{aligned} \text{div } \mathbf{H} &= 0, \\ \text{div } \mathbf{D} &= 0. \end{aligned} \right\} \dots\dots\dots (2.4)$$

\* See Stratton's "Electromagnetic Theory". (McGraw-Hill, New York : 1941.)

The equations (2.1), (2.2), and (2.3) determine the electromagnetic field in a magneto-ionic medium completely. Since they are linear and homogeneous in all field vectors it is sufficient to consider only harmonic time-variations. A more complex time-variation may then be obtained by a Fourier summation of harmonic components. Accordingly we take the field components to be the real parts of

$$\left. \begin{aligned} \mathbf{E} &= \mathbf{E}' e^{-ipt}, \\ \mathbf{H} &= \mathbf{H}' e^{-ipt}, \text{ etc.,} \end{aligned} \right\} \dots\dots\dots (2.5)$$

where  $\mathbf{E}'$ ,  $\mathbf{H}'$ , etc. are functions of position only.

The equations (2.1) and (2.3) then become

$$\left. \begin{aligned} \text{curl } \mathbf{E} - ip\mu_v \mathbf{H} &= 0, \\ \text{curl } \mathbf{H} + ip\mathbf{D} &= 0, \end{aligned} \right\} \dots\dots\dots (2.6)$$

and

$$-p(p + iv)\mathbf{P} + ip \frac{e\mu_v}{m} \mathbf{P} \times \mathbf{H}_0 = \frac{Ne^2}{m} \mathbf{E} \dots\dots\dots (2.7)$$

The components of  $\mathbf{P}$  are therefore related to those of  $\mathbf{E}$  by the components of a complex susceptibility matrix  $(\sigma_{ij})$  such that

$$P_i = \epsilon_v \sum_{j=1}^3 \sigma_{ij} E_j \dots\dots\dots (2.8)$$

If the coordinate system is rectangular we can write down the components of the inverse matrix  $(\sigma_{ij})^{-1} = (\sigma_{ij}^{(-1)})$ , say, from (2.7). Then

$$x(\sigma_{ij}^{(-1)}) = \begin{pmatrix} -1 - iz & iy_3 & -iy_2 \\ -iy_3 & -1 - iz & iy_1 \\ iy_2 & -iy_1 & -1 - iz \end{pmatrix} \dots\dots (2.9)$$

Here we have introduced the very convenient dimensionless quantities of Appleton(12)

$$x = \frac{p_0^2}{p^2}, \quad p_0^2 = \frac{Ne^2}{m\epsilon_v} \dots\dots\dots (2.10)$$

where  $p_0/2\pi$  is the natural frequency with which the electron plasma constituting the medium will oscillate in the absence of a magnetic field if the collision frequency is negligible,

$$y = \frac{p_H}{p}, \quad p_H = \frac{e\mu_v H_0}{m} \dots\dots\dots (2.11)$$

where  $p_H/2\pi$ , the magnitude of the vector  $p_H/2\pi$ , is the gyro-frequency of electron motion in the magnetic field  $\mathbf{H}_0$ ,  $y$  the magnitude of  $y$  and

$$z = \frac{v}{p} \dots\dots\dots (2.12)$$

The matrix  $(\sigma_{ij})$  is equivalent to Hartree's scattering tensor(1) or Saha and Banerjee's conductivity tensor(7).

The relation between the components of  $\mathbf{D}$  and  $\mathbf{E}$  can similarly be expressed by the components of a complex dielectric matrix  $(K_{ij})$ , using (2.2) and (2.8).

$$D_i = \epsilon_v \sum_{j=1}^3 K_{ij} E_j \dots\dots\dots (2.13)$$

where

$$\left. \begin{aligned} K_{ij} &= \sigma_{ij} + \delta_{ij}, \\ \delta_{ij} &= \begin{cases} 1 & \text{if } i=j, \\ 0 & \text{if } i \neq j. \end{cases} \end{aligned} \right\} \dots\dots\dots (2.14)$$

Eliminating  $\mathbf{H}$  in (2.6) the condition on  $\mathbf{E}$  becomes

$$\text{curl curl } \mathbf{E} - p^2 \mu_v \mathbf{D} = 0 \dots\dots\dots (2.15)$$

When  $\mathbf{E}$  has been determined from (2.15), (2.13), and the initial and boundary conditions of the problem,  $\mathbf{H}$  can be found from the first relation (2.6).

However, if these conditions involve  $\mathbf{H}$  as well as  $\mathbf{E}$  or if the source distribution of the field is specified, it is more convenient to have both  $\mathbf{E}$  and  $\mathbf{H}$  dependent on one auxiliary function. The most appropriate potential is the electric Hertz vector  $\mathbf{Z}$ , whose source is equivalent to a distribution of charge and current, quite arbitrary save that it must satisfy the usual continuity condition. The specification of initial and boundary conditions on  $\mathbf{Z}$  ensures that the corresponding values of  $\mathbf{E}$  and  $\mathbf{H}$  satisfy the field conditions. If

$$\left. \begin{aligned} \mathbf{E} &= \text{grad div } \mathbf{Z} + p^2 \mu_v \epsilon_v \mathbf{Z}, \\ \mathbf{H} &= -ip \epsilon_v \text{curl } \mathbf{Z}, \end{aligned} \right\} \dots\dots\dots (2.16)$$

the field equations impose on  $\mathbf{Z}$  the condition

$$\text{curl curl } \mathbf{Z} - \frac{1}{\epsilon_v} \mathbf{D} = 0 \dots\dots\dots (2.17)$$

where  $\mathbf{D}$  is related to  $\mathbf{Z}$  through (2.13) and (2.16).

We could also express the field in terms of the magnetic Hertz vector  $\mathbf{Z}^*$ . Such fields, however, are only particular cases of  $\mathbf{Z}$ -type fields, as can be seen most readily by writing

$$\mathbf{Z} = -\frac{1}{ip \epsilon_v} \text{curl } \mathbf{Z}'$$

when a  $\mathbf{Z}^*$ -type field results. The same relationship holds for any linear electromagnetic field. The source of a  $\mathbf{Z}^*$ -type field is equivalent only to distributions for which the charge density is zero, whence the current density distribution must be solenoidal. It appears that the advantages of using "complex" Hertz vectors of the type  $\mathbf{Z} \mp i\mathbf{Z}^*$ , as suggested by Bateman(13) and others, are largely nebulous.

### III. THE PRINCIPAL DIRECTIONS OF THE MEDIUM

It is a usual simplification in the investigation of phenomena in anisotropic media to use a coordinate system in which the axes are parallel to the principal directions in the medium, if these be all different. In a magneto-ionic medium we seek the directions in which  $\mathbf{P}$  is parallel to  $\mathbf{E}$ . Then  $\mathbf{D}$  will also be parallel to  $\mathbf{E}$ . We must then have, in terms of the original components,

$$\begin{aligned} P_i &= \epsilon_v \sum_{j=1}^3 \sigma_{ij} E_j \\ &= \epsilon_v \sigma E_i, \end{aligned} \dots\dots\dots (3.1)$$

say. It is, however, simpler to write (3.1) in the inverse form

$$\left. \begin{aligned} \epsilon_v E_i &= \sum_{j=1}^3 \sigma_{ij}^{(-1)} P_j \\ &= \sigma^{-1} P_i, \end{aligned} \right\} \dots\dots\dots (3.2)$$

where the components of  $(\sigma_{ij}^{(-1)})$  are given by (2.9).

Then  $\sigma$  is such that

$$\sum_{j=1}^3 (\sigma_{ij}^{(-1)} - \sigma^{-1} \delta_{ij}) P_j = 0, \dots\dots\dots (3.3)$$

and for these three equations to be consistent we must have

$$|| \sigma_{ij}^{(-1)} - \sigma^{-1} \delta_{ij} || = 0 \dots\dots\dots (3.4)$$

This equation yields three values for  $\sigma$  which satisfy (3.1) or (3.2) and the corresponding directions may be found from any two of the equations (3.3). These values are of course the eigenvalues of the matrix, and the corresponding directions its eigenvectors. Using the rectangular coordinate system we find the eigenvalues are

$$\left. \begin{aligned} \sigma_1 &= -\frac{x}{1+iz-y}, \\ \sigma_2 &= -\frac{x}{1+iz+y}, \\ \sigma_3 &= -\frac{x}{1+iz}. \end{aligned} \right\} \dots\dots\dots (3.5)$$

Saha and Banerjea(7) obtained this result but did not pursue the implications as far as we do. We find that the principal directions, or eigenvectors, are parallel to  $\mathbf{P}$  when its rectangular components are in the ratios

$$\left. \begin{aligned} \frac{P_1}{y_2^2 + y_3^2} &= \frac{P_2}{-y_1 y_3 - i y y_3} = \frac{P_3}{-y_3 y_1 + i y y_2}, \\ \frac{P_1}{y_2^2 + y_3^2} &= \frac{P_2}{-y_1 y_3 + i y y_3} = \frac{P_3}{-y_3 y_1 - i y y_2}, \\ \frac{P_1}{y_1} &= \frac{P_2}{y_2} = \frac{P_3}{y_3}, \end{aligned} \right\} \dots\dots\dots (3.6)$$

corresponding respectively to  $\sigma_1$ ,  $\sigma_2$ , and  $\sigma_3$ . The first two eigenvectors belong to the complex number field. They are complex conjugates and are therefore not orthogonal. The third is real and can be shown to be orthogonal to both of the others. It is parallel to  $\mathbf{y}$  and this suggests choosing our "physical" coordinate system so that one of the axes has the direction of  $\mathbf{H}_0$ . For any rectilinear system this requires that the direction of  $\mathbf{H}_0$  be invariant throughout the medium. An appropriate "physical" coordinate system is the rectangular system  $O\mathbf{x}_1\mathbf{x}_2\mathbf{x}_3$  where the direction of  $O\mathbf{x}_3$  is parallel to  $\mathbf{H}_0$ . Then

$$\mathbf{y} = (0, 0, -y) \dots\dots\dots (3.7)$$

since the charge on the electron is negative, and the eigenvectors are such that

$$\left. \begin{aligned} \frac{P_1}{1} &= \frac{P_2}{i} = \frac{P_3}{0}, \\ \frac{P_1}{1} &= \frac{P_2}{-i} = \frac{P_3}{0}, \\ \frac{P_1}{0} &= \frac{P_2}{0} = \frac{P_3}{1}. \end{aligned} \right\} \dots\dots\dots (3.8)$$

We now introduce a "complex" rectilinear coordinate system whose axes are parallel to the eigenvectors of the medium. The relations between the components of  $\mathbf{E}$ ,  $\mathbf{P}$ , and  $\mathbf{D}$  and the conditions they must satisfy then take their



simplest forms. Coordinates in the physical system are converted into coordinates in the fundamental "mathematical" system  $O\mathbf{x}_1\mathbf{x}_2\mathbf{x}_3$ , whose axes are parallel to the eigenvectors, by means of the linear transformation

$$\left. \begin{aligned} \mathbf{x}_1 &= \mathbf{x}_1 + \mathbf{x}_2, & \mathbf{x}_1 &= \frac{1}{2}(x_1 - ix_2), \\ \mathbf{x}_2 &= i(\mathbf{x}_1 - \mathbf{x}_2), & \mathbf{x}_2 &= \frac{1}{2}(x_1 + ix_2), \\ \mathbf{x}_3 &= \mathbf{x}_3, & \mathbf{x}_3 &= x_3. \end{aligned} \right\} \dots\dots (3.9)$$

We shall denote contravariant components of  $\mathbf{E}$ ,  $\mathbf{H}$ ,  $\mathbf{P}$ ,  $\mathbf{D}$ , and  $\mathbf{Z}$  in the fundamental mathematical system by  $\mathbf{E}_i$ ,  $\mathbf{H}_i$ ,  $\mathbf{P}_i$ ,  $\mathbf{D}_i$ , and  $\mathbf{Z}_i$ . Then we have

$$\left. \begin{aligned} \mathbf{P}_1 &= \epsilon_v \sigma_1 \mathbf{E}_1, \\ \mathbf{P}_2 &= \epsilon_v \sigma_2 \mathbf{E}_2, \\ \mathbf{P}_3 &= \epsilon_v \sigma_3 \mathbf{E}_3, \end{aligned} \right\} \dots\dots\dots (3.10)$$

and

$$\left. \begin{aligned} \mathbf{D}_1 &= \epsilon_v K_1 \mathbf{E}_1, \\ \mathbf{D}_2 &= \epsilon_v K_2 \mathbf{E}_2, \\ \mathbf{D}_3 &= \epsilon_v K_3 \mathbf{E}_3, \end{aligned} \right\} \dots\dots\dots (3.11)$$

where from (2.14) and (3.5)

$$\left. \begin{aligned} K_1 &= 1 - \frac{x}{1 + iz - y}, \\ K_2 &= 1 - \frac{x}{1 + iz + y}, \\ K_3 &= 1 - \frac{x}{1 + iz}. \end{aligned} \right\} \dots\dots\dots (3.12)$$

The eigenvalues  $\sigma_1$ ,  $\sigma_2$ , and  $\sigma_3$  of  $(\sigma_{ij})$  are the principal complex susceptibilities and  $K_1$ ,  $K_2$ , and  $K_3$  are the principal complex dielectric constants of the magneto-ionic medium. Writing

$$\left. \begin{aligned} K_1 &= q_1^2, \\ K_2 &= q_2^2, \\ K_3 &= q_3^2, \end{aligned} \right\} \dots\dots\dots (3.13)$$

we recognize  $q_1$  and  $q_2$  as the complex refractive indices of the two plane waves which can be propagated parallel to  $\mathbf{H}_0$  in a uniform medium, and  $q_3$  as the complex refractive index of the "ordinary" wave of transverse propagation. The refractive index  $q_4$  of the "extraordinary" wave of transverse propagation is evidently not a fundamental quality in the theory. If we write

$$\begin{aligned} K_4 &= q_4^2 \\ &= 1 - \frac{x(1 + iz - x)}{(1 + iz)(1 + iz - x) - y^2} \dots\dots\dots (3.14) \end{aligned}$$

we find that it is a function of  $K_1$  and  $K_2$ ,

$$K_4 = \frac{2K_1 K_2}{K_1 + K_2} \dots\dots\dots (3.15)$$

To determine the significance of the first two mathematical components,  $\mathbf{E}_1$  and  $\mathbf{E}_2$  say, we consider the complex Argand representation\* of the physical field components  $R\mathbf{E}_1$  and  $R\mathbf{E}_2$ . If we denote complex conjugates by an asterisk, we have

$$\begin{aligned} R\mathbf{E}_1 + iR\mathbf{E}_2 &= \frac{1}{2}(\mathbf{E}_1 + i\mathbf{E}_2) + \frac{1}{2}(\mathbf{E}_1^* + i\mathbf{E}_2^*) \\ &= \mathbf{E}_2 + \mathbf{E}_1^* \dots\dots\dots (3.16) \end{aligned}$$

\* The author is indebted to Dr. E. O. Hercus, Physics Department, University of Melbourne, for this analysis.

Now if  
as in (2.5), we can write

$$p = \omega - i\alpha, \quad \dots\dots\dots (3.17)$$

$$\begin{aligned} \mathbf{E}_2 &= \mathbf{E}'_2 e^{-ipt} \\ &= |\mathbf{E}'_2| e^{-\alpha t - i(\omega t - \beta_2)}, \quad \dots\dots\dots (3.18) \end{aligned}$$

where

$$\cos \beta_2 = \frac{Rl \mathbf{E}'_2}{|\mathbf{E}'_2|}, \quad \sin \beta_2 = \frac{Im \mathbf{E}'_2}{|\mathbf{E}'_2|} \quad \dots\dots\dots (3.19)$$

This represents a damped rotation, with angular velocity  $\omega$ , of the field component perpendicular to  $Ox_3$  in a negative (left-handed) sense about  $Ox_3$ .

Similarly,

$$\begin{aligned} \mathbf{E}_1^* &= \mathbf{E}'_1^* e^{ip^*t} \\ &= |\mathbf{E}'_1| e^{-\alpha t + i(\omega t - \beta_1)} \quad \dots\dots\dots (3.20) \end{aligned}$$

where

$$\cos \beta_1 = \frac{Rl \mathbf{E}'_1}{|\mathbf{E}'_1|}, \quad \sin \beta_1 = \frac{Im \mathbf{E}'_1}{|\mathbf{E}'_1|} \quad \dots\dots\dots (3.21)$$

This represents a damped rotation in the positive (right-handed) sense about  $Ox_3$ .

Hence generally the field component perpendicular to the imposed magnetic field  $\mathbf{H}_0$  may be analysed into two opposite rotating fields, of equal angular velocity and damping, about the direction of  $\mathbf{H}_0$ .

#### IV. THE EQUATIONS FOR $\mathbf{E}$ AND $\mathbf{Z}$ IN THE FUNDAMENTAL COORDINATE SYSTEM

In order to make use of the fundamental coordinate system  $O\mathbf{x}_1\mathbf{x}_2\mathbf{x}_3$  we must find the representations of the standard vector operations. For this purpose we note that the contravariant components of vectors in the mathematical system are derived from the components in the physical system by a transformation similar to (3.9).\*

Then if  $\psi$  and  $\mathbf{F}$  be arbitrary functions of position we have

$$\begin{aligned} \text{grad}_1 \psi &= \frac{1}{2}(\text{grad}_1 - i \text{grad}_2)\psi \\ &= \frac{1}{2}\left(\frac{\partial}{\partial x_1} - i \frac{\partial}{\partial x_2}\right)\psi \\ &\quad - \frac{1}{2} \frac{\partial \psi}{\partial \mathbf{x}_3} \end{aligned}$$

by (3.9). The other results follow in a similar manner. Grouping them we have (in all cases the components are contravariant)

$$\left. \begin{aligned} \text{grad}_1 \psi &= \frac{1}{2} \frac{\partial \psi}{\partial \mathbf{x}_2}, \\ \text{grad}_2 \psi &= \frac{1}{2} \frac{\partial \psi}{\partial \mathbf{x}_1}, \\ \text{grad}_3 \psi &= \frac{\partial \psi}{\partial \mathbf{x}_3}. \end{aligned} \right\} \dots\dots\dots (4.1)$$

\* See McConnell's "Applications of the Absolute Differential Calculus". Ch. II, Section 2. (Blackie & Son Ltd., 1931.)

$$\operatorname{div} \mathbf{F} = \frac{\partial \mathbf{F}_1}{\partial \mathbf{x}_1} + \frac{\partial \mathbf{F}_2}{\partial \mathbf{x}_2} + \frac{\partial \mathbf{F}_3}{\partial \mathbf{x}_3} \dots \dots \dots (4.2)$$

$$\left. \begin{aligned} \operatorname{curl}_1 \mathbf{F} &= -i \left( \frac{\partial \mathbf{F}_1}{\partial \mathbf{x}_3} - \frac{1}{2} \frac{\partial \mathbf{F}_3}{\partial \mathbf{x}_2} \right), \\ \operatorname{curl}_2 \mathbf{F} &= i \left( \frac{\partial \mathbf{F}_2}{\partial \mathbf{x}_3} - \frac{1}{2} \frac{\partial \mathbf{F}_3}{\partial \mathbf{x}_1} \right), \\ \operatorname{curl}_3 \mathbf{F} &= i \left( \frac{\partial \mathbf{F}_1}{\partial \mathbf{x}_1} - \frac{\partial \mathbf{F}_2}{\partial \mathbf{x}_2} \right). \end{aligned} \right\} \dots \dots \dots (4.3)$$

$$\nabla^2 \psi = \operatorname{div} \operatorname{grad} \psi = \frac{\partial^2 \psi}{\partial \mathbf{x}_1 \partial \mathbf{x}_2} + \frac{\partial^2 \psi}{\partial \mathbf{x}_3^2} \dots \dots \dots (4.4)$$

It follows from these results that

$$\left. \begin{aligned} (\operatorname{grad}_1 \operatorname{div} - \operatorname{curl}_1 \operatorname{curl}) \mathbf{F} &= \nabla^2 \mathbf{F}_1, \\ (\operatorname{grad}_2 \operatorname{div} - \operatorname{curl}_2 \operatorname{curl}) \mathbf{F} &= \nabla^2 \mathbf{F}_2, \\ (\operatorname{grad}_3 \operatorname{div} - \operatorname{curl}_3 \operatorname{curl}) \mathbf{F} &= \nabla^2 \mathbf{F}_3, \end{aligned} \right\} \dots \dots \dots (4.5)$$

as is the case for components in a rectangular system.

Now from these results and (3.11) we can write down the fundamental component equations of (2.15),

$$\left. \begin{aligned} \left( \nabla^2 + \frac{p^2}{c^2} K_1 \right) \mathbf{E}_1 - \frac{1}{2} \frac{\partial}{\partial \mathbf{x}_2} \operatorname{div} \mathbf{E} &= 0, \\ \left( \nabla^2 + \frac{p^2}{c^2} K_2 \right) \mathbf{E}_2 - \frac{1}{2} \frac{\partial}{\partial \mathbf{x}_1} \operatorname{div} \mathbf{E} &= 0, \\ \left( \nabla^2 + \frac{p^2}{c^2} K_3 \right) \mathbf{E}_3 - \frac{\partial}{\partial \mathbf{x}_3} \operatorname{div} \mathbf{E} &= 0, \end{aligned} \right\} \dots \dots \dots (4.6)$$

where  $c$  is the velocity of electromagnetic waves in free space,

$$c^2 = \frac{1}{\mu_v \epsilon_v} \dots \dots \dots (4.7)$$

These equations do not agree with those of Saha *et al.* (7, 9), where the term in  $\operatorname{div} \mathbf{E}$  in each equation is omitted.

Again we can write down the fundamental component equations in  $\mathbf{Z}$  of (2.17),

$$\left. \begin{aligned} \left( \nabla^2 + \frac{p^2}{c^2} K_1 \right) \mathbf{Z}_1 + \frac{1}{2} (K_1 - 1) \frac{\partial}{\partial \mathbf{x}_2} \operatorname{div} \mathbf{Z} &= 0, \\ \left( \nabla^2 + \frac{p^2}{c^2} K_2 \right) \mathbf{Z}_2 + \frac{1}{2} (K_2 - 1) \frac{\partial}{\partial \mathbf{x}_1} \operatorname{div} \mathbf{Z} &= 0, \\ \left( \nabla^2 + \frac{p^2}{c^2} K_3 \right) \mathbf{Z}_3 + (K_3 - 1) \frac{\partial}{\partial \mathbf{x}_3} \operatorname{div} \mathbf{Z} &= 0. \end{aligned} \right\} \dots \dots (4.8)$$

The equations (4.6) and (4.8) in the mathematical coordinate system again display the fundamental nature of the principal dielectric constants of the medium. They are much simpler in form than the corresponding equations in the physical system.

The conditions on the components of  $\mathbf{E}$  and  $\mathbf{Z}$  are not in general independent. If  $\operatorname{div} \mathbf{E} = 0$  the fields corresponding to each component are independent and may

be considered separately. For a uniform medium, elementary wave-solutions are of the form

$$\mathbf{E} = \mathbf{E}'' e^{i(\mathbf{k}_1 \mathbf{x}_1 + \mathbf{k}_2 \mathbf{x}_2 + \mathbf{k}_3 \mathbf{x}_3 - pt)},$$

where

$$\mathbf{k}_1 \mathbf{k}_2 + \mathbf{k}_3^2 = \frac{p^2}{c^2} \bar{K} \dots \dots \dots (4.9)$$

with appropriate subscripts on  $\mathbf{E}$ ,  $\mathbf{E}''$ , and  $\bar{K}$ . Any three of  $\mathbf{k}_1$ ,  $\mathbf{k}_2$ ,  $\mathbf{k}_3$ , and  $p$  may be chosen arbitrarily. In the physical coordinate system (4.9) is equivalent to the plane wave-solution

$$\mathbf{E} = \mathbf{E}'' e^{i(k_1 x_1 + k_2 x_2 + k_3 x_3 - pt)},$$

where

$$k_1^2 + k_2^2 + k_3^2 = \frac{p^2}{c^2} K, \dots \dots \dots (4.10)$$

and again three of  $k_1$ ,  $k_2$ ,  $k_3$ , and  $p$  are arbitrary. The same may be said for the components of  $\mathbf{Z}$  under the condition  $\text{div } \mathbf{Z} = 0$ .

Another simple case is when the field depends only on the coordinate  $\mathbf{x}_3$ . Then (4.6) and (4.8) give

$$\left. \begin{aligned} \left( \frac{\partial^2}{\partial \mathbf{x}_3^2} + \frac{p^2}{c^2} K_1 \right) \mathbf{E}_1 &= 0, \\ \left( \frac{\partial^2}{\partial \mathbf{x}_3^2} + \frac{p^2}{c^2} K_2 \right) \mathbf{E}_2 &= 0, \\ \mathbf{E}_3 &= 0, \end{aligned} \right\} \dots \dots \dots (4.11)$$

and

$$\left. \begin{aligned} \left( \frac{\partial^2}{\partial \mathbf{x}_3^2} + \frac{p^2}{c^2} K_1 \right) \mathbf{Z}_1 &= 0, \\ \left( \frac{\partial^2}{\partial \mathbf{x}_3^2} + \frac{p^2}{c^2} K_2 \right) \mathbf{Z}_2 &= 0, \\ \left( \frac{\partial^2}{\partial \mathbf{x}_3^2} + \frac{p^2}{c^2} \right) \mathbf{Z}_3 &= 0. \end{aligned} \right\} \dots \dots \dots (4.12)$$

In both cases the component fields are independent, with those associated with  $\mathbf{E}_3$  and  $\mathbf{Z}_3$  zero. The elementary wave-solutions for the remaining components in a uniform medium are of the form

$$\left. \begin{aligned} \mathbf{E} &= \mathbf{E}'' e^{i(\kappa \mathbf{x}_3 - pt)}, \\ \mathbf{Z} &= \mathbf{Z}'' e^{i(\kappa \mathbf{x}_3 - pt)}, \\ \mathbf{k}^2 &= \frac{p^2}{c^2} \bar{K}. \end{aligned} \right\} \dots \dots \dots (4.13)$$

where

With appropriate subscripts, these are the familiar plane-wave solutions of longitudinal propagation, that is, of propagation parallel to  $\mathbf{H}_0$ .

Now if the field be transverse, there is no dependence on  $\mathbf{x}_3$  and we find

$$\left. \begin{aligned} \left( \frac{\partial^2}{\partial \mathbf{x}_1 \partial \mathbf{x}_2} + \frac{p^2}{c^2} K_1 \right) \mathbf{E}_1 - \frac{1}{2} \frac{\partial}{\partial \mathbf{x}_2} \left( \frac{\partial \mathbf{E}_1}{\partial \mathbf{x}_1} + \frac{\partial \mathbf{E}_2}{\partial \mathbf{x}_2} \right) &= 0, \\ \left( \frac{\partial^2}{\partial \mathbf{x}_1 \partial \mathbf{x}_2} + \frac{p^2}{c^2} K_2 \right) \mathbf{E}_2 - \frac{1}{2} \frac{\partial}{\partial \mathbf{x}_1} \left( \frac{\partial \mathbf{E}_1}{\partial \mathbf{x}_1} + \frac{\partial \mathbf{E}_2}{\partial \mathbf{x}_2} \right) &= 0, \\ \left( \frac{\partial^2}{\partial \mathbf{x}_1 \partial \mathbf{x}_2} + \frac{p^2}{c^2} K_3 \right) \mathbf{E}_3 &= 0, \end{aligned} \right\} \dots \dots \dots (4.14)$$



and

$$\left. \begin{aligned} \left( \frac{\partial^2}{\partial \mathbf{x}_1 \partial \mathbf{x}_2} + \frac{p^2}{c^2} K_1 \right) \mathbf{Z}_1 + \frac{1}{2} (K_1 - 1) \frac{\partial}{\partial \mathbf{x}_2} \left( \frac{\partial \mathbf{Z}_1}{\partial \mathbf{x}_1} + \frac{\partial \mathbf{Z}_2}{\partial \mathbf{x}_2} \right) &= 0, \\ \left( \frac{\partial^2}{\partial \mathbf{x}_1 \partial \mathbf{x}_2} + \frac{p^2}{c^2} K_2 \right) \mathbf{Z}_2 + \frac{1}{2} (K_2 - 1) \frac{\partial}{\partial \mathbf{x}_1} \left( \frac{\partial \mathbf{Z}_1}{\partial \mathbf{x}_1} + \frac{\partial \mathbf{Z}_2}{\partial \mathbf{x}_2} \right) &= 0, \\ \left( \frac{\partial^2}{\partial \mathbf{x}_1 \partial \mathbf{x}_2} + \frac{p^2}{c^2} K_3 \right) \mathbf{Z}_3 &= 0. \end{aligned} \right\} \dots (4.15)$$

Only the fields associated with  $\mathbf{E}_3$  and  $\mathbf{Z}_3$  are independent. For a uniform medium, elementary wave-solutions are of the form

$$\left. \begin{aligned} \mathbf{E}_3 &= \mathbf{E}_3'' e^{i(\kappa_1 \mathbf{x}_1 + \kappa_2 \mathbf{x}_2 - p t)} \\ \mathbf{Z}_3 &= \mathbf{Z}_3'' e^{i(\kappa_1 \mathbf{x}_1 + \kappa_2 \mathbf{x}_2 - p t)}, \\ \mathbf{k}_1 \mathbf{k}_2 &= \frac{p^2}{c^2} K_3 \end{aligned} \right\} \dots (4.16)$$

where

and two of  $\mathbf{k}_1$ ,  $\mathbf{k}_2$ , and  $p$  are arbitrary. This is the familiar "ordinary" plane-wave solution of transverse propagation. The other "extraordinary" wave is evidently compounded from  $\mathbf{E}_1$ ,  $\mathbf{E}_2$  or  $\mathbf{Z}_1$ ,  $\mathbf{Z}_2$  components. If we assume these components have the same form as (4.16), (4.14) and (4.15) give two linear homogeneous conditions on each pair,

$$\left. \begin{aligned} \left( \frac{1}{2} \mathbf{k}_1 \mathbf{k}_2 - \frac{p^2}{c^2} K_1 \right) \mathbf{E}_1 - \frac{1}{2} \mathbf{k}_2^2 \mathbf{E}_2 &= 0, \\ \frac{1}{2} \mathbf{k}_1^2 \mathbf{E}_1 - \left( \frac{1}{2} \mathbf{k}_1 \mathbf{k}_2 - \frac{p^2}{c^2} K_2 \right) \mathbf{E}_2 &= 0, \end{aligned} \right\} \dots (4.17)$$

and

$$\left. \begin{aligned} \left( \frac{1}{2} (K_1 + 1) \mathbf{k}_1 \mathbf{k}_2 - \frac{p^2}{c^2} K_1 \right) \mathbf{Z}_1 + \frac{1}{2} (K_1 - 1) \mathbf{k}_2^2 \mathbf{Z}_2 &= 0, \\ \frac{1}{2} (K_2 - 1) \mathbf{k}_1^2 \mathbf{Z}_1 + \left( \frac{1}{2} (K_2 + 1) \mathbf{k}_1 \mathbf{k}_2 - \frac{p^2}{c^2} K_2 \right) \mathbf{Z}_2 &= 0. \end{aligned} \right\} \dots (4.18)$$

The first pair is consistent only under the condition

$$\begin{aligned} \mathbf{k}_1 \mathbf{k}_2 &= \frac{p^2}{c^2} \frac{2K_1 K_2}{K_1 + K_2}, \\ &= \frac{p^2}{c^2} K_4 \dots (4.19) \end{aligned}$$

by (3.15); the second pair is consistent under this condition and also if

$$\mathbf{k}_1 \mathbf{k}_2 = p^2 / c^2,$$

but since then the field due to  $\mathbf{Z}_1$  and  $\mathbf{Z}_2$  is zero, only the condition (4.19) is relevant. This is the result we should expect. The "extraordinary" wave of transverse propagation then consists of components  $\mathbf{E}_1$ ,  $\mathbf{E}_2$  or  $\mathbf{Z}_1$ ,  $\mathbf{Z}_2$  in the proportions given by either of the equations in (4.17) or (4.18), where they depend on the direction of propagation.

## V. PLANE WAVES

In the Introduction we saw that plane waves can be propagated in any direction in a magneto-ionic medium only if the properties of the medium are

uniform. Under these conditions we suppose that the direction of propagation is parallel to a unit vector  $\mathbf{n}$  whose components in the physical coordinate system are given in terms of spherical polar angles  $\theta, \varphi$ ,

$$\mathbf{n} = (\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta) \dots \dots \dots (5.1)$$

We are not likely to require the  $\mathbf{Z}$ -representation of this case and therefore consider only the set of equations (4.6). Then  $\mathbf{E}$  will have the form

$$\begin{aligned} \mathbf{E} &= \mathbf{E}'' \exp \left[ -ip \left\{ t - \frac{q}{c} (x_1 \sin \theta \cos \varphi + x_2 \sin \theta \sin \varphi + x_3 \cos \theta) \right\} \right] \\ &= \mathbf{E}'' \exp \left[ -ip \left\{ t - \frac{q}{c} (\mathbf{x}_1 e^{i\varphi} \sin \theta + \mathbf{x}_2 e^{-i\varphi} \sin \theta + \mathbf{x}_3 \cos \theta) \right\} \right] \dots (5.2) \end{aligned}$$

by (3.9). The complex refractive index  $q$  and the direction of  $\mathbf{E}''$  are determined by  $\mathbf{n}$  and the electromagnetic properties of the medium. Substitution into (4.6) yields three linear homogeneous conditions on the components,

$$\left. \begin{aligned} \{ K(1 + \cos^2 \theta) - 2K_1 \} \mathbf{E}_1 - K e^{-2i\varphi} \sin^2 \theta \mathbf{E}_2 - K e^{-i\varphi} \cos \theta \sin \theta \mathbf{E}_3 &= 0, \\ -K e^{2i\varphi} \sin^2 \theta \mathbf{E}_1 + \{ K(1 + \cos^2 \theta) - 2K_2 \} \mathbf{E}_2 - K e^{i\varphi} \cos \theta \sin \theta \mathbf{E}_3 &= 0, \\ -K e^{i\varphi} \sin \theta \cos \theta \mathbf{E}_1 - K e^{-i\varphi} \sin \theta \cos \theta \mathbf{E}_2 + (K \sin^2 \theta - K_3) \mathbf{E}_3 &= 0, \end{aligned} \right\} \dots \dots \dots (5.3)$$

where we have introduced a new function

$$K = q^2, \dots \dots \dots (5.4)$$

which we may call the *associated complex dielectric constant* of the wave in the medium.  $K$  is determined by the condition that the three equations (5.3) should be consistent. Proceeding as in Section III, we find that this requires

$$\begin{aligned} K^2 \{ (K_1 + K_2) \sin^2 \theta + 2K_3 \cos^2 \theta \} - K \{ 2K_1 K_2 \sin^2 \theta + (K_1 - K_2) K_3 (1 + \cos^2 \theta) \} \\ + 2K_1 K_2 K_3 = 0 \dots \dots \dots (5.5) \end{aligned}$$

Using the relation (3.15) we obtain the alternative form

$$\begin{aligned} K^2 (K_1 K_2 \sin^2 \theta + K_3 K_4 \cos^2 \theta) - K K_1 K_2 \{ K_3 (1 + \cos^2 \theta) + K_4 \sin^2 \theta \} \\ + K_1 K_2 K_3 K_4 = 0 \dots \dots \dots (5.6) \end{aligned}$$

Substitution in (5.5) of the values (3.12) leads to the well-known Appleton-Hartree formula for  $K = q^2$ . The forms (5.5) and (5.6) are very suggestive and we make extensive use of them in a paper on classical magneto-ionic theory to be published shortly. The proportions of  $\mathbf{E}_1, \mathbf{E}_2, \mathbf{E}_3$  in either of the two plane waves for which  $K$  satisfies (5.5) or (5.6) can be found from any two of the equations (5.3). The result is

$$\begin{aligned} \frac{\mathbf{E}_1}{K(K - K_2)e^{-i\varphi} \cos \theta \sin \theta} &= \frac{\mathbf{E}_2}{K(K - K_1)e^{-i\varphi} \cos \theta \sin \theta} \\ &= \frac{\mathbf{E}_3}{2K^2 \cos^2 \theta - K(K_1 + K_2)(1 + \cos^2 \theta) + 2K_1 K_2} \dots \dots \dots (5.7) \end{aligned}$$

## VI. PLASMA OSCILLATIONS

The characteristic of a pure-oscillatory field is that its variation with time is independent of position. Inspection of (4.6) and (4.8) shows that this can happen only under special conditions. In particular, the medium should be uniform within its boundaries. We shall restrict our consideration to the equations (4.6) in  $\mathbf{E}$ . Then there are three possible types of oscillatory field

associated with  $\mathbf{E}_1$ ,  $\mathbf{E}_2$ , and  $\mathbf{E}_3$  respectively. The condition for the first type is  $K_1 = \mathbf{E}_2 = \mathbf{E}_3 = 0$ . The characteristics of frequency and damping are given by  $K_1 = 0$  whence

$$x = 1 + iz - y \quad \dots\dots\dots (6.1)$$

and

$$p = -\frac{1}{2}(iv - p_H) \pm \sqrt{\{p_0^2 + \frac{1}{4}(iv - p_H)^2\}}$$

when the values (2.10), (2.11), and (2.12) are substituted for  $x$ ,  $y$ , and  $z$ . If

$$p = \omega - i\alpha \quad \dots\dots\dots (6.2)$$

$\alpha$  is positive and the alternative must be chosen such that  $\omega$  is positive. Then

$$p = p_1 = -\frac{1}{2}(iv - p_H) + \sqrt{\{p_0^2 + \frac{1}{4}(iv - p_H)^2\}} \quad \dots\dots\dots (6.3)$$

Since  $\alpha$  is zero when  $v$  is neglected, we see that the effect of collisions is to damp the oscillation. If  $v$  is small compared with  $p_0$  and  $p_H$  we can approximate to (6.3) to get

$$\left. \begin{aligned} \omega &= \omega_1 \doteq \sqrt{(p_0^2 + \frac{1}{4}p_H^2) + \frac{1}{2}p_H} \\ \alpha &= \alpha_1 \doteq \frac{1}{2}v \left\{ 1 + \frac{\frac{1}{2}p_H}{\sqrt{(p_0^2 + \frac{1}{4}p_H^2)}} \right\} \end{aligned} \right\} \quad \dots\dots\dots (6.4)$$

The condition on  $\mathbf{E}_1$  is

$$\left( \frac{1}{2} \frac{\partial^2}{\partial \mathbf{x}_1 \partial \mathbf{x}_2} + \frac{\partial^2}{\partial \mathbf{x}_3^2} \right) \mathbf{E}_1 = 0 \quad \dots\dots\dots (6.5)$$

The type of oscillation executed by  $\mathbf{E}_1$  can be found by considering the corresponding real physical components. We have

$$\left. \begin{aligned} \mathbf{E}_1 &= \mathbf{E}_1' e^{i\omega_1 t - \alpha_1 t}, \\ \mathbf{E}_2 &= \mathbf{E}_3 = 0, \end{aligned} \right\} \quad \dots\dots\dots (6.6)$$

whence from (3.9) and (3.16)

$$\left. \begin{aligned} R\mathbf{E}_1 + iR\mathbf{E}_2 &= \mathbf{E}_1^* \\ R\mathbf{E}_3 &= 0 \end{aligned} \right\} \quad \dots\dots\dots (6.7)$$

Thus the real electric vector rotates in a positive (right-handed) sense about the direction of  $\mathbf{H}_0$ , with angular velocity  $\omega_1$  and damping constant  $\alpha_1$ , as given by (6.2) and (6.3). The other field vectors do likewise.

The second type of oscillation is associated with  $\mathbf{E}_2$  when  $K_2 = \mathbf{E}_3 = \mathbf{E}_1 = 0$ . Corresponding to (6.1), (6.3), (6.4), and (6.5) we find

$$x = 1 + iz + y \quad \dots\dots\dots (6.9)$$

$$p = p_2 = -\frac{1}{2}(iv + p_H) + \sqrt{\{p_0^2 + \frac{1}{4}(iv + p_H)^2\}} \quad \dots\dots (6.10)$$

$$\omega = \omega_2 \doteq \sqrt{(p_0^2 + \frac{1}{4}p_H^2) - \frac{1}{2}p_H} \quad \dots\dots\dots (6.11)$$

$$\alpha = \alpha_2 \doteq \frac{1}{2}v \left\{ 1 - \frac{\frac{1}{2}p_H}{\sqrt{(p_0^2 + \frac{1}{4}p_H^2)}} \right\}, \quad \dots\dots\dots (6.12)$$

and

$$\left( \frac{1}{2} \frac{\partial^2}{\partial \mathbf{x}_1 \partial \mathbf{x}_2} + \frac{\partial^2}{\partial \mathbf{x}_3^2} \right) \mathbf{E}_2 = 0 \quad \dots\dots\dots (6.13)$$

Proceeding as before we find that the field rotates about the direction of  $\mathbf{H}_0$  in a negative (left-handed) sense, with angular velocity  $\omega_2$  and damping constant  $\alpha_2$ .

Finally, for the third type of oscillation, associated with  $\mathbf{E}_3$  when

$$K_3 = \mathbf{E}_1 = \mathbf{E}_2 = 0,$$

we find

$$x = 1 + iz \dots\dots\dots (6.14)$$

$$\omega = \omega_3 \doteq p_0 \dots\dots\dots (6.15)$$

$$\alpha = \alpha_3 \doteq \frac{1}{2}\nu \dots\dots\dots (6.16)$$

and

$$\frac{\partial^2 \mathbf{E}_3}{\partial \mathbf{x}_1 \partial \mathbf{x}_3} = 0 \dots\dots\dots (6.17)$$

The oscillation is linear in the direction of  $\mathbf{H}_0$  with frequency  $\omega_3/2\pi$  and damping constant  $\alpha_3$ .

It appears that there are three possible types of oscillation of electron plasma of given density and collision frequency in the presence of a magnetic field, having the characteristics just investigated. Tonks and Langmuir(14) showed that the frequency  $p_0/2\pi$  was the natural frequency of electron plasma in the absence of a magnetic field, neglecting collisions.

Martyn(15) reported that in the presence of a magnetic field there are two natural frequencies of oscillation of electron plasma. He suggested that such oscillations in the solar atmosphere might be responsible for the emission of solar radio-frequency noise of intensity greater than the "quiet" level.

He stated further that the effect of collisions in an ionized medium is to alter the oscillatory state of the electromagnetic field (which implies no propagation) to a state where radiation is propagated. He suggested that the source of this radiation might lie in the borders of prominence material reaching into the corona, where the higher density and lower temperature, corresponding to chromospheric conditions, should enhance the radiating power of the medium (presumably because of the higher collision frequency). The oscillations could be maintained by the breakdown in electric polarization due to turbulent eddies which occur there.

It is apparent from the above analysis that there are, in fact, three types of oscillation of a magneto-ionic medium. The effect of collisions is not to alter the oscillatory character of the electromagnetic field but to damp the oscillations and to reduce the natural frequencies. The very high collision frequency in chromospheric material will effectively damp out any oscillation altogether.

The process of radiation from a plasma oscillation depends, as in any radiating aerial system, on a difference in electromagnetic properties between oscillator and the surrounding medium. The propagation constant of a uniform oscillating medium is zero; any non uniformity such as a gradient or discontinuity in electron density or magnetic field intensity will lead to a non-zero propagation constant for the frequency of oscillation. Radiation of the plasma frequency will therefore be propagated from the boundaries of the oscillating medium. It appears that some mechanism such as Martyn's is required to maintain the oscillation.

This discussion highlights the difference between the two roles of collisions in radiation from a macroscopic plasma oscillation of an ionized medium and in thermal radiation due to the individual microscopic processes occurring within



the medium. The thermal emission per unit volume and time is proportional to the frequency of electron-ion collisions and the radiation emitted during each collision (Smerd and Westfold 16). The plasma oscillation is superimposed in an ordered manner on all these microscopic processes. Collisions, while contributing individually to the thermal emissivity, tend to damp-out the macroscopic radiation process whose electromagnetic field can propagate only from the boundaries.

It is clear that local plasma oscillations taking place in the solar atmosphere can result in the propagation of radiation of frequencies  $\omega_1$ ,  $\omega_2$ , or  $\omega_3$ . Associated with these are the damping constants  $\alpha_1$ ,  $\alpha_2$ , and  $\alpha_3$  which are of the order of  $\frac{1}{4}\nu$  to  $\frac{3}{4}\nu$  under such conditions. The characteristics of frequency and damping will be maintained during propagation. However, the refractive index structure of the solar atmosphere is such that it is unlikely that radiation of the second type will escape.

Ruby Payne-Scott(17) of this Laboratory has recently made an analysis of the characteristics of a particular class of high-level "bursts" of solar noise. These exhibit a more or less exponential decay, with damping constants of the order of the collision frequency  $\nu$  at the level in the corona where  $\omega_1$ ,  $\omega_2$ , or  $\omega_3$  equals the frequency on which a burst is received.

## VII. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Radiophysics, C.S.I.R.O. The author is indebted to Dr. J. C. Jaeger, University of Tasmania, and Dr. E. O. Hercus, University of Melbourne, for much helpful advice and criticism.

## VIII. REFERENCES

- (1) HARTREE, D. R.—*Proc. Camb. Phil. Soc.* **27**: 143 (1931).
- (2) BOOKER, H. G.—*Proc. Roy. Soc. A* **155**: 235 (1936).
- (3) APPLETON, E. V.—*Rep. Un. Radio-Sci. Inst. Washington* (1927).
- (4) NICHOLS, H. W., and SCHELLING, J. C.—*Bell Syst. Tech. J.* **4**: 215 (1925).
- (5) BAKER, W. G., and GREEN, A. L.—*Proc. Instn. Radio Engrs.* **21**: 1103 (1933).
- (6) HUXLEY, L. G. H.—*Phil. Mag.* **25**: 148 (1938).
- (7) SAHA, M. N., and BANERJEA, B. K.—*Indian J. Phys.* **19**: 159 (1945).
- (8) BANERJEA, B. K.—*Proc. Roy. Soc. A* **190**: 67 (1947).
- (9) SAHA, M. N., BANERJEA, B. K., and GUHA, U. C.—*Indian J. Phys.* **21**: 181 (1947).
- (10) BAILEY, V. A.—*Proc. Roy. Soc. N.S.W.* **82**: 107 (1948); *Aust. J. Sci. Res. A* **1**: 351 (1948).
- (11) DARWIN, C. G.—*Proc. Roy. Soc. A* **146**: 17 (1934).
- (12) APPLETON, E. V.—*J. Inst. Elect. Engng.* **71**: 642 (1932).
- (13) BATEMAN, H.—"Electrical and Optical Wave-motion." (Cambridge Univ. Press: Cambridge, 1915.)
- (14) TONKS, L., and LANGMUIR, I.—*Phys. Rev.* **33**: 195 (1929).
- (15) MARTYN, D. F.—*Nature* **159**: 26 (1947).
- (16) SMERD, S. F., and WESTFOLD, K. C.—*Phil. Mag.* (in press).
- (17) PAYNE-SCOTT, R.—*Aust. J. Sci. Res. A* **2**: 214.

# THE STRUCTURE OF COSMIC RAY AIR SHOWERS

By J. R. PRESCOTT\* and C. B. O. MOHR†

[*Manuscript received March 11, 1949*]

## *Summary*

The structure of cosmic ray air showers at sea-level has been studied by an investigation of the burst rate frequency and the transition effect in lead, for cosmic ray bursts occurring simultaneously in two high-pressure ionization chambers with varying separation.

Although extensive showers were responsible for all the coincidences observed with the larger chamber separations, they accounted for less than 3 per cent. of the bursts observed with a single chamber. Of the remaining 97 per cent., somewhat more than one-half appear to be due to nuclear disintegrations and the rest either to narrow showers of approximate radius 30 cm. or to the core of an extensive shower of low density. The extensive shower frequency was about 10 times that predicted by theory.

The bearing of these results on present views of the origin and development of air showers is discussed.

## I. INTRODUCTION

The discovery of cosmic ray air showers, often called extensive showers, was reported in 1938(1, 2). They were shown to cover large areas and to be capable of causing coincident discharges in Geiger counters several hundred metres apart.

To explain these showers, the cascade theory of shower production was extended by taking into account the angular scattering of the shower particles. Starting from the assumption that electrons of very high energy are present near the top of our atmosphere, cascade theory describes the multiplication of such particles into many photons and electrons by successive radiative collisions and pair production, the energy being divided among the secondary particles produced (see, for example, Heisenberg 3). As a result of multiple scattering processes, the secondary particles are spread out laterally in the course of their passage through the atmosphere to form extensive showers. The form of the distribution function obeyed by the shower electrons has been investigated by Molière(4); it predicts a shower with a highly condensed core of high energy electrons, and an outer fringe of electrons with lower energy.

The theory has been applied(1) to calculate the rate of discharge of pairs of coincidence counters as a function of counter separation, but experiments(5) have shown a much more rapid increase in the coincidence rate with diminishing

\* Atomic Physics Section, C.S.I.R.O., Physics Department, University of Melbourne.

† Physics Department, University of Melbourne.

separation than calculated. Geiger and Stubbe(6), recording fivefold and sixfold coincidences, also found an unexpectedly high counting rate at small counter separations. Alichanian and Asatiani(7) obtained a similar result with improved counter systems and they attributed the effect to narrow showers, i.e. cascade showers having their origin well below the top of the atmosphere and having a relatively small spread. Molière suggested, however, that at small separations the counters detect cores of air showers of low initial energy which are well past their maximum development, and for which existing cascade theory calculations are invalid.

The observed rate of occurrence of air showers at low altitudes(8) and its increase with altitude(9), agree reasonably well with theory. In these experiments the counter separations were sufficiently large to avoid the effect just described, and thus ensure that only extensive showers were being recorded.

However, experiments with counters give an incomplete picture, for counters do not give directly the density of the shower, and results obtained with their use have to be compared with the results of theoretical calculations in which the estimated probability of discharge of the counter must be averaged over all sizes of shower. There is therefore a great advantage in carrying out experiments with a pair of coincidence ionization chambers, as these give directly the density at two points in the shower. However, the few such experiments which have been carried out give a picture that is still not complete.

Lewis(10), observing at 3100 metres altitude with two thin-walled ionization chambers containing argon at 1.3 atmospheres pressure, found that as the chamber separation was decreased below 5 metres, the coincidence rate increased much more rapidly than predicted by theory for extensive showers (see Fig. 4). A little later, Kingshill and Lewis(11), using the same chambers at sea-level with a separation of 1 metre, found that only about 2 per cent. of the bursts occurring in one chamber were accompanied by a burst in the other. Both these sets of results were attributed to local showers.

Recently Bridge *et al.*(12), observing at 3500 metres with four thin-walled chambers containing argon at 5 atmospheres pressure, also found a sharp increase in the coincidence rate when their chambers were brought within a few inches of each other. By analysing the shape of their pulses, and also by using a Wilson cloud chamber fired by impulses from an ionization chamber built into it, they were able to show that most single chamber bursts were caused by heavily ionizing particles from nuclear disintegrations in the walls of the chamber, initiated by cosmic ray particles. They therefore attributed most of their chamber coincidences to nuclear disintegrations, and Bethe(13) suggested that nuclear disintegrations were also responsible for the results of Kingshill and Lewis. The same might equally well have been said about the results of Lewis for his smaller chamber separations.

Williams(14), using the ionization chambers of Bridge *et al.*(12) at altitudes greater than 3000 metres, found that the coincidence rate showed no very marked increase for chamber separations down to 0.15 metre (see Fig. 4).

Theory(15, 16) gives a coincidence rate for ionization chambers recording showers of a given density, which rises very slowly with decreasing separation



and which does not show any sharp rise at small separations; though a more recent treatment of the problem by Blatt(17), which has not been published in detail, appears to predict a very sharp increase in coincidence rate for chamber separations less than about 1 metre.

In the present work, which was carried out at sea-level with two high-pressure chambers, unshielded, and also shielded with different thicknesses of lead, curves were obtained for the coincidence rate as a function of chamber separation. For such curves we shall adopt the convenient term *decoherence curve*, used by Williams.

## II. APPARATUS

### (a) The Chambers

The ionization chambers used in this work were made from drawn steel tubing of inner diameter 9.7 cm. and wall thickness 0.175 cm.; their construction is shown in Figure 1. They contained nitrogen at a pressure of 39 atmospheres.

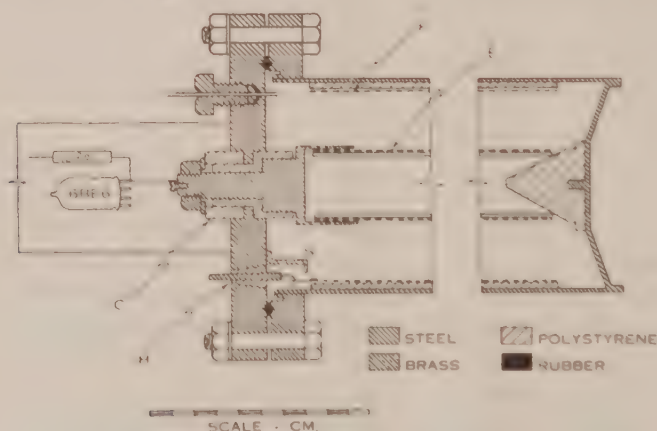


Fig. 1.—Cross section of ionization chamber.

The two electrodes, *E*, *F*, in Figure 1, which were of copper gauze, were 3.3 cm. and 8.8 cm. in diameter and 58 cm. in length. The outer electrode was held away from the chamber wall by six polystyrene strips. At the gas pressure used,  $\alpha$  particles from radioactivity of the walls could not therefore penetrate into the effective volume of the chamber; constructing the electrodes of gauze reduces ionization due to their natural radioactivity.

Connections to both electrodes were brought out through individual polystyrene plugs *C*, *H* in the end flange of the chamber. With this construction the flange, which is held at earth potential, acts as a guard ring between the two electrodes. The inner (collecting) electrode was prevented from "seeing" any of the surface of the high tension insulator by a brass shield *S*. A positive potential of 400 volts was applied to the outer electrode. This potential was supplied by Exide 10-volt high tension cells, each mounted as an individually insulated unit.



*(b) The Amplifiers*

The inner electrode of each chamber was connected directly to the control grid of a 6BE6 valve. This valve when used with appropriately reduced voltages on the electrodes, acted as a combined electrometer valve and preamplifier of voltage gain about 50(18). The valve was specially treated to improve insulation between the pins(18), and was mounted in an air-tight metal box containing a drying agent, the box itself being screwed to the flange of the ionization chamber.

To allow the ionization due to general background radiation to leak away, the input grid was earthed through a resistance of  $10^{12}$  ohms. The measured capacity of the chamber and input grid combined was  $38 \mu\mu\text{F}$ .

The preamplifier was followed by two stages of direct coupled amplification with negative feedback to the first grid of the 6BE6, giving an overall voltage gain with feedback of about 400. The voltage output of the amplifier was then applied through a  $2 \mu\text{F}$  condenser to the grid of a magic eye tuning indicator (6U5/6G5). The whole equipment was battery operated, and all units and connecting leads were carefully shielded.

*(c) The Recording System*

Images of one edge of the shadow on the target of each 6U5 were thrown in juxtaposition on a defining slit of width 0.0018 cm. by lenses of focal length 3.4 cm. and aperture  $f:3.4$ . The movement of the shadow edge in response to the occurrence of a burst in the chamber was recorded on 35 mm. film moving at a speed of 2 mm. per minute at right angles to the slit and in contact with it.

The resolving time for coincidences was of the same order as the time of collection of the ions, which was 0.4 second. For measuring purposes the film was enlarged five times by projection.

Magic eyes were first used by Montgomery and Montgomery(19) in portable ionization chamber equipment, where the common method of recording the deflection of a sensitive galvanometer could not conveniently be used. This method has the big advantage of eliminating zero drift, and the recording system has a negligible time of response. It has the disadvantage that the screens of the magic eyes fatigue and become much less bright with continued use, necessitating the replacement of the valves after a very few weeks.

The apparatus was calibrated by finding the deflections produced on the film when a series of very small graded changes was made in the high potential applied to the outer electrode of the chamber. Taking for the specific ionization in nitrogen 70 ion pairs per cm.\*; for the mean path length of a ray through the chamber the value 5.8 cm., and for the degree of saturation 63 per cent.(23, 24), it was found that the smallest kick on the film observable above the statistical background corresponded to a burst in the chamber of 30 particles, a mean shower density of 600 particles per square metre, and an energy loss of 16 MeV.

\* The specific ionization of shower electrons of energy 100 MeV. was determined from the cloud chamber observations of Corson and Brode(20), Sen Gupta(21), and Hazen(22). Allowance was made for the difference in efficiency of cloud chambers and ionization chambers for detecting ionization.

Observations were made on the frequency of occurrence of bursts of this size and greater in either of the two chambers.

The chambers themselves were placed with axes horizontal and mutually parallel, on light, wooden platforms resting on the rafters of a galvanized iron hut with an asbestos roof of thickness  $1.5 \text{ g./cm.}^2$ . The distance to the nearest building (of one storey) was 40 ft., and other buildings lay at much greater distances.

To study the transition effect in lead, and also to utilize the lead as an "intensifier" of the somewhat infrequent coincidences, lead sheets were bent round the top half of the chambers into the form of hemicylindrical caps.

### III. DISCUSSION OF RESULTS

#### (a) *The Decoherence Curves*

The coincidence rates observed with different separations of the chamber axes and with different thicknesses of lead above the chambers are shown by the points on the diagram in Figure 2. Curves A, B, and C correspond to runs in which the chambers were covered with 0, 0.5, and 1 cm. of lead respectively; the ordinates of curves B and C have been multiplied by 0.18 and 0.1 respectively to make these curves coincide with curve A for the larger chamber separations. A logarithmic scale is used for the vertical axis, and a square-law scale for the horizontal axis in order to be able to show the value 0; the values shown for *zero separation* of the chambers are the burst rates for a *single* chamber. The probable form of the curves through the experimental values is shown by the full lines; the limits of error indicated for each value refer to the standard error calculated from the number of coincidences. These limits are rather large and the data somewhat incomplete (the measurements were brought to an end through the destruction of the apparatus by fire); nevertheless, there is sufficient information for certain general conclusions to be drawn. The figures represent 1600 hours' observations taken between December 1947 and September 1948.

It will be seen from Figure 2 that only a small fraction of the events causing a burst in one chamber also cause an observable burst in the other at the same instant, even when the chambers are only 16 cm. apart; also there are indications that the coincidence rate is beginning to increase as the separation is decreased through these smaller values, whereas for the larger separations the coincidence rate alters comparatively little.

Theoretical calculations show that atmospheric showers will have a very large lateral extension, and a relatively small dense core(1). This, however, will so infrequently pass near the chambers that the coincidence rate for showers having locally the observed density (600 particles per square metre) will vary relatively slowly with chamber separation, even for small separations(16) (see dotted curve in Fig. 2). The coincidences observed for chamber separations greater than 32 cm. are undoubtedly due to extensive showers, but any large increase at the smaller separations (including the greatly increased burst rate in a single chamber) must be due to some other effect, viz. nuclear disintegrations or local showers or both. These conclusions may have to be modified in the light of the calculations of Blatt(17), mentioned earlier.

*(b) Nuclear Disintegrations*

It was first pointed out by Euler(15) that some of the smaller bursts observed in ionization chambers might be due to a few highly ionizing particles (from nuclear disintegrations) instead of a much larger number of lightly ionizing particles (shower electrons), and Carmichael(25) has recently confirmed this in a detailed analysis of experimental data. Montgomery and Montgomery(26) endeavoured to separate bursts due to nuclear disintegrations from those due to showers by analysis of the burst frequency—burst size distributions obtained

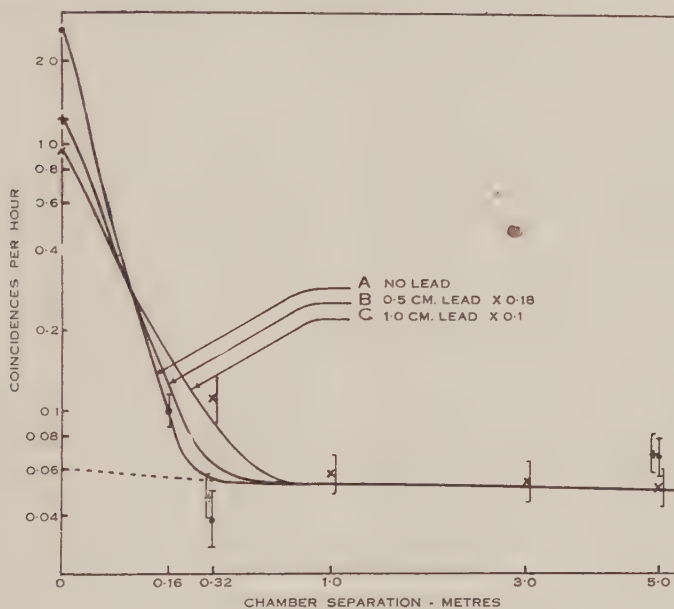


Fig. 2.—Coincidence rates for bursts of more than 30 particles, corresponding to a mean shower density of more than 600 particles. The curves for coincidence rates with lead above the chambers are adjusted to coincide with the curve for no lead at the larger chamber separations, by the use of appropriate multiplying factors.

with different thicknesses of burst-producing material above the chamber. Bridge *et al.*(12) showed that most of the bursts in a single chamber were due to particles from nuclear disintegrations, and that many of the particles were able to penetrate thicknesses of metal greater than that of the chamber walls. They concluded that at least the greater part of their greatly increased coincidence rate for closely packed chambers must have been due to nuclear disintegrations.

That many of our single chamber bursts were also due to nuclear disintegrations was shown in a few preliminary runs taken prior to the fire, mentioned earlier, when discharges of a fourfold coincidence counter system placed immediately above one chamber were registered on the film along with the bursts in that chamber. Most of the bursts were unaccompanied by a discharge of the counters, showing that they must have been produced, not by showers,



but by nuclear disintegrations produced in the walls of the chamber by a single particle, which would have been unable to discharge the coincidence counters. Insufficient data were obtained, however, to give a reliable value for the proportion of the bursts produced by showers only.

The increased coincidence rate for a chamber separation of 16 cm. may also be partly or wholly due to these nuclear disintegrations. A rough calculation, based on the solid angle subtended at one chamber by the other, shows that the number of coincidences due to disintegration particles emitted in all directions from the walls of the chamber will fall to a value comparable with the number due to extensive showers, when the separation has increased to 0.3 metre, assuming that all the disintegration particles have sufficient energy to penetrate the walls of the chambers.

There are difficulties, however, in explaining how sufficient ionization for a 16 MeV. burst can be produced in one chamber by disintegration protons from the other chamber. The greatest number of ions will be obtained from a proton with just sufficient energy to cross the chamber near a diametral plane, viz. from a proton still having an energy of 17 MeV. after penetrating into the chamber; but the fraction of the ions collected from such a proton track for our field and gas pressure will be less than half that for an electron track (25), so that the proton would produce a burst of less than 8 MeV. At least three slow protons would therefore be required to produce all but the very smallest bursts measured. For protons with greater energy than is required just to traverse the chamber, the energy loss in the gas would be less, and still more protons would be required to produce the observed bursts. Furthermore, it was found that surrounding one chamber with a copper cylinder of wall thickness 0.57 cm., when the separation was 0.16 metre, produced no appreciable reduction of the coincidence rate, and each proton would require at least 75 MeV. energy to pass through two chamber walls and the copper cylinder and still have sufficient energy to contribute to a burst.

While protons of such high energies are found in "stars" or nuclear disintegrations in photographic plates—Bridge *et al.* (12) also observed protons with energies above 60 MeV.—most protons have energies of the order of 10 MeV., and only a very small proportion possess these higher energies (27, 28); and very seldom would several of these protons be found within the relatively small solid angle subtended by one chamber at the other. If many of the disintegration particles contributing to bursts are  $\alpha$  particles, as shown by Carmichael, their energies would have to be considerably higher. Again some coincidences might be ascribed to nuclear disintegrations produced in the wall of one chamber by mesons or neutrons from nuclear disintegrations in the wall of the other chamber, but the probability of such penetrating particles ending their life in the chamber wall would be remotely small.

It therefore seems more feasible to attribute the increased coincidence rate at the smaller separations to local showers, where we use the term "local shower" to indicate a region of high particle density of limited extent. Such showers are either the cores of "old" extensive showers well past their maximum or



"young" showers still in the course of development. Evidence in support of these conclusions will be given in Section III (d).

(c) *Local Showers: The Transition Curves*

Let us consider the information given by the transition curves obtained for lead. Figure 3 shows the coincidence frequency as a function of lead thickness above both chambers for chamber separations of 5 and 0.32 metres; and the single chamber frequency as a function of lead thickness. The ordinates of the two former curves have been multiplied by 20 for comparative purposes. The curve for 0.32 metre separation is steeper at the larger thicknesses than that for 5 metres separation, which is understandable on the assumption of local showers;

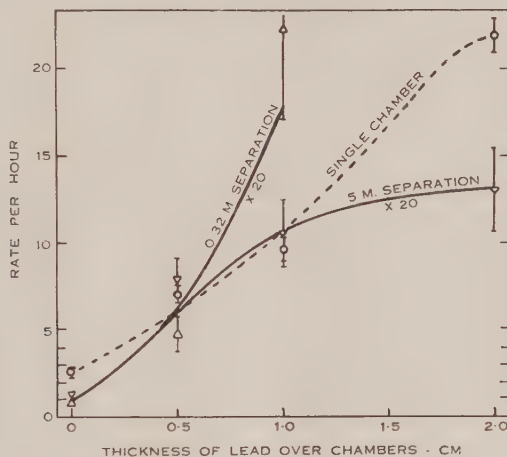


Fig. 3.—Transition curve for lead. The ordinates of the curves for 0.32 metre and 5 metres separation have been multiplied by 20 for comparative purposes.

for the coincidences at the larger separation will be due to shower electrons with energies not greatly exceeding 100 MeV.(16), the critical energy for air, while those at the smaller separation arise partly or wholly from local showers, which have been shown by counter experiments(29) to be more penetrating and which will therefore multiply more in the lead.\*

The single chamber rate might thus be expected to be still steeper (the curves being adjusted to the same height at one point, viz. 0.5 cm. Pb, in Fig. 3), but for the following consideration. If coincident bursts were always of the same size in both chambers, placing lead above both would increase the burst size equally in both, and hence increase the coincidence rate by the same factor that it would increase the rate in either chamber alone. Analytical treatment

\* It seems unlikely that sufficient coincidences could be caused by widely diverging shower electrons from the lead over the chambers(30), though it is conceivable that a contributing factor might be nuclear disintegrations in the lead.

shows that if the bursts are of somewhat different sizes in the two chambers, placing lead above both chambers will increase the coincidence rate *relatively* more than it will increase the rate of either chamber alone; the greater the disparity between the burst sizes, the greater will be the relative increase. As the bursts in each chamber are in fact of different sizes, owing to differences of density within the shower, it is understandable that the curve for the single chamber should be less steep than the curve for 0.32 metre separation.

This argument can be made plausible by an intuitive picture. We draw a diagram in which each coincidence is represented by a point in the quadrant enclosed by the positive  $x$ - and  $y$ -axes, the size of the burst in each chamber being represented by the distance of the point from the  $x$ - and  $y$ -axes respectively. The points will be mainly clustered about the line  $x=y$ , the amount of scatter to either side of this line being determined by Molière's structure function for showers(4). At the same time we must satisfy the empirical relation that the number of bursts in a single chamber greater than a given size  $s$  must vary as  $s^{-\gamma}$ , where  $\gamma$  is about 1.5. The number of coincidences in which the burst size in both chambers is greater than  $A$  is found by counting the points in the region  $x > A, y > A$ . In practice,  $A$  corresponds to the size of the smallest burst detectable in the chambers. Placing lead above the chambers multiplies the number of rays in the chamber and makes smaller bursts detectable, and this has the effect of lowering the value of  $A$  to a value say,  $a$ . Thus the number of bursts in one chamber is now the number of points in the region  $x > a$  or  $y > a$ , whereas the number of coincidences is the number of points in the region  $x > a, y > a$ . More points are added in the second case than in the first, i.e. the relative increase in coincidence rate is greater than the relative increase in the single chamber rate. The analytical treatment, which is not simple, consists in expressing this picture in mathematical terms.

#### (d) Local Showers: Analysis of the Single Chamber Burst Rate

The curves in Figure 3 for 5 and 0.32 metres separation correspond to bursts produced by showers only, while the curve for the single chamber also includes bursts produced by nuclear disintegrations. The latter will not be affected notably in frequency by the addition of the lead(31), and will represent merely a constant amount to be subtracted from each point of the single chamber curve to give a curve for the transition effect for showers only. If we extrapolate the flat portion of curve  $A$  in Figure 2 in accordance with the theoretical calculations of Wolfenstein, we obtain a rate of 0.06 per hour for bursts due to extensive showers. If we now assume that the balance of bursts in a single chamber without lead shielding is due to nuclear disintegrations alone, we obtain a rate for disintegrations of 2.54 per hour, which is practically constant for all thicknesses of lead. Subtracting this value from the single chamber rate of 6.2 per hour under 0.5 cm. Pb we obtain a shower rate of 3.7 per hour. This represents a 60-fold increase over the rate with no lead, an increase which is impossibly large. One is forced to conclude that disintegrations cannot account for such a large proportion of the bursts in a single chamber.

The work of Montgomery and Montgomery with a single chamber, in which nuclear disintegrations are allowed for in deducing transition curves for showers only, leads to ratios of 1:4.3:8.6 for the relative shower frequencies with 0, 0.5, and 1 cm. Pb respectively, for showers of the same density as ours. By fitting our transition curve to theirs, we obtain a figure of 1.45 per hour for nuclear disintegrations in our chamber. Thus if this figure is subtracted from the single chamber transition curve, we get rates of 1.1, 4.8, and 9.3 per hour for the three thicknesses, and these rates are in the ratio 1:4.4:8.5, in close agreement with the results of the Montgomerys just quoted.\*

For two chambers at 5 metres separation, the coincidence rates at 0, 0.5, and 1 cm. Pb are in the ratio 1:6:10.5, as compared with the ratio 1:4.4:8.5 for a single chamber. The more rapid increase with thickness for this coincidence rate is to be expected in the light of the discussion in Section III (c).

The burst rate of 2.6 per hour for a single chamber is therefore compounded of a rate for nuclear disintegrations of about 1.45 per hour, and a rate for showers of about 1.1 per hour. This shower frequency of 1.1 per hour includes an extensive shower frequency of 0.06 per hour, so that the great majority of showers seem to be local showers. The probable form of the decoherence curve allowing for nuclear disintegrations is shown in Figure 4 (curve D).

We may at this point comment on a feature of Figure 2. When the curves for different thicknesses of lead are arbitrarily adjusted to coincide at the larger chamber separations, the points for zero separation fall successively lower for increasing thicknesses of lead. This is due to the fact that while showers are multiplied by the lead, disintegrations, which have been shown to contribute a large fraction of the bursts at zero separation, are little affected by the presence of lead.

#### (e) Comparison with Other Observers

In view of these results, let us now consider briefly the results of Bridge *et al.*(12) and of Lewis(10) at 3500 and 3100 metres respectively. The effect of the increased altitude will be to increase the frequency of nuclear disintegrations by a slightly greater factor than the extensive showers(32). Now consider the effect of gas pressure. Our chambers contained nitrogen at 39 atmospheres pressure, while those of Bridge *et al.*—of the same shape and very similar dimensions—contained argon at 5 atmospheres pressure. The much smaller pressure will decrease the ionization due to shower particles, while leaving unaltered that due to disintegration particles whose tracks end in the chamber (as will be the case for the most frequently occurring disintegrations). The smaller low energy limit for bursts detected by Bridge *et al.* also favours the detection of nuclear disintegrations at the expense of showers. Another factor affecting the detection of disintegrations is the nature of the gas in the chambers. In argon, recombination of ions is very low even for tracks of highly ionizing particles, whereas

\* It is possible to think of burst-producing mechanisms other than local showers, which will give an increased burst rate when lead is placed above the chamber, e.g. the production of nuclear disintegrations by photons in the photonuclear effect, but it is difficult to estimate the contribution of such an effect.



in nitrogen recombination of the ions in the tracks of disintegration particles is very much greater than it is in the tracks of the much more thinly ionizing shower particles(25). The contribution of showers in the experiments of Bridge *et al.* will therefore be very much less, and that of nuclear disintegrations correspondingly greater, than in our experiments.

Lewis's chambers were thin-walled, contained argon at 1.3 atmospheres pressure, and had a relatively large volume, being spheres of diameter 35 cm. The small pressure was therefore compensated for by the longer path length in the chamber, and the equivalent path length at normal pressure was about the same as in chambers of Bridge *et al.*, so that nuclear disintegrations would not be expected to play a much larger part than in experiments of Bridge *et al.*, in spite of the lower pressure. Lewis's coincidence frequencies rose more rapidly than would be expected for extensive showers for separations less than about 3 metres (see Fig. 4), although the contribution from nuclear disintegrations should not have become greater than the contribution from extensive showers until the separation was 0.4 metre or less, according to a calculation based on the solid angle subtended by one chamber at the other. It might thus seem that local showers were contributing appreciably to the coincidence rate for separations up to 2 or 3 metres, in Lewis's experiments at 3100 metres. On the other hand, Williams(14) at a similar altitude finds no such increase in frequency even down to 0.15 metre. While one feels that reliance must be placed on the results of Williams, it is difficult to account for the different form of curve obtained by Lewis at comparatively large separations. Williams states that he finds no evidence for local showers.

In the light of these remarks, it is of interest to consider the results of the counter experiments of Alichanian and Asatiani(7) on narrow showers, in which possible coincidences due to particles produced in the material of the counter walls were eliminated by the counter arrangement. A narrow shower radius of about 3 metres was obtained at an altitude of 3250 metres as against a radius of about 0.4 metre at an altitude of 960 metres, the same counter arrangement being used in both cases. Furthermore, the use of a counter system favouring the selection of denser showers was found to give a greater value for the radius of the shower; this might be expected, as the denser showers require a higher energy primary to produce them.

#### (f) *Extensive Showers*

A quantity of importance for theory is the frequency of occurrence of extensive showers of given density, and the variation of frequency with altitude. To be certain that one is observing only extensive showers one must use coincidence chambers sufficiently far apart. The present work leads to a single chamber rate of 0.06 per hour for showers of density greater than 600 particles per square metre at sea-level, a value which is 10 times larger than Wolfenstein's calculated value of 0.006 per hour for showers of the same density.

As the coincidence rate observed at 3100 metres by Lewis for his largest chamber separation of 10 metres must have been due only to extensive showers, his value of 7 per hour for showers of the same density as ours is 280 times



Wolfenstein's(16) calculated value of 0.025 per hour for a chamber separation of 10 metres and for an altitude of 3100 metres. Even the considerably smaller frequency of 0.3 per hour, obtained from Williams's results at 3050 metres by reducing his showers to the same density as ours, is 12 times Wolfenstein's value. The discrepancy with theory thus appears to become even more marked with increase in altitude.

The relation of the observed to the calculated results at the two altitudes is shown more clearly in Figure 4, where the calculated values of Euler for sea-level and for the same shower density are also added for comparative purposes. The calculations of Wolfenstein are an improvement on those of Euler(15), since they are based on a more accurate value of the spread of the shower and

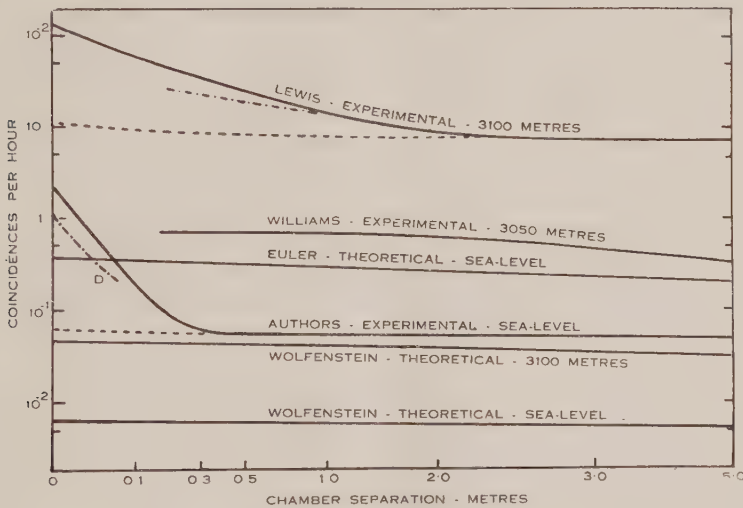


Fig. 4.—Comparison of experimental and theoretical curves for showers of average density greater than 600 particles per square metre.

- Extensive showers.
- . - . - . Extensive showers + local showers (probable form of curve).
- Extensive showers + local showers + nuclear disintegrations, for the experimental curves.

The theoretical curves are for extensive showers.

also take into account the zenith effect, neglected by Euler. The very large difference in the results of the two calculations indicates a great sensitivity of the results to the various factors involved, and suggests that further improvements in the calculations, e.g. a more accurate estimation of the number of electrons with energies below the critical energy, might well lead to different results again. Recent work(14) seems to indicate that in the lower atmosphere the direction of incidence of nearly all extensive showers is close to the vertical, in which case the zenith angle effect will have been overestimated in Wolfenstein's calculations, and part—but by no means all—of the discrepancy is accounted for. The calculations of Blatt(17), already mentioned, may give better agreement, but no details are at present available.

*(g) Origin of Air Showers*

While, for the reasons given, the theoretical calculations may not yet be regarded as final, and while the absolute magnitude of the experimental values may be somewhat inaccurate, it seems most unlikely that the remarkable discrepancy between theory and experiment can be wholly accounted for in this way.

There seems no alternative but to adopt Wolfenstein's suggestion\* that the large frequencies observed for extensive showers are due to these showers originating, not just at the top of the atmosphere as is usually thought, but also well below it.

The further assumption of narrow showers, some of them originating still lower in the atmosphere and not becoming fully developed, would indicate that air showers are produced throughout a considerable depth of the atmosphere.

In view of what has been said, the agreement between experimental and theoretical shower frequencies (8, 9) for showers of *lower* density as measured with *counters* is extremely puzzling. A point that requires further investigation is whether the local showers, whose existence has been demonstrated in this paper, are cores of extensive showers well past their maximum or narrow showers produced in the lower atmosphere.

Further experiments with new apparatus are being carried out to clarify this point and to confirm the previous results. A discussion of the burst size-frequency distributions and of the transition curves at larger thicknesses of lead will form the subject of a subsequent communication.

## IV. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Atomic Physics Section, C.S.I.R.O. We wish to thank Professor L. H. Martin for his continued interest in this work.

## V. REFERENCES

- (1) AUGER, P., MAZE, R., and GRIVET-MEYER, T.—*C.R. Acad. Sci. Paris* **206** : 1721 (1938).
- (2) SCHMEISER, K., and BOTHE, W.—*Ann. Phys. Lpz.* **32** : 161 (1938).
- (3) HEISENBERG, W.—“Cosmic Radiation.” Ch. II. (Ed. by W. Heisenberg, Dover Publications: New York, 1946.)
- (4) MOLIERE, G.—*Ibid.* Ch. III.

\* *Note Added in Proof.*—We are indebted to Dr. R. Williams for suggesting a more likely explanation. Wolfenstein used an energy spectrum of the form  $E^{-\gamma}$  with  $\gamma=1.8$ , which he assumed to hold up to energies of  $10^{16}$  eV. Changing  $\gamma$  from 1.8 to 1.63 would increase the calculated values by a factor of 10, and bring them into reasonable agreement with experiment.

A personal communication from Blatt indicates that the theoretical decoherence curve rises by a factor of about two between two metres and zero separation at 3100 metres altitude. This would still not account for the big increase observed in the present experiments. Blatt also points out that in “old” showers the core is very much less pronounced. It appears, therefore, that the observed increase in coincidence rate at small separations is due to narrow showers still in the course of development.

We are indebted to Dr. J. M. Blatt for communicating his results to us in advance of publication.

- (5) AUGER, P., MAZE, R., and ROBLEY, J.—*C.R. Acad. Sci. Paris* **208** : 1641 (1939).
- (6) GEIGER, H., and STUBBE, W.—*Abh. preuss. Akad. Wiss.* **3** (10) : (1941).
- (7) ALICHANIAN, A., and ASATIANI, T.—*J. Phys. U.S.S.R.* **9** : 175 (1945).
- (8) COCCONI, G., LOVERDO, A., and TONGIORGI, V.—*Phys. Rev.* **70** : 846 (1946).
- (9) HILBERRY, N.—*Ibid.* **60** : 1 (1941).
- (10) LEWIS, L. G.—*Ibid.* **67** : 228 (1945).
- (11) KINGSHILL, K. L., and LEWIS, L. G.—*Ibid.* **69** : 159 (1946).
- (12) BRIDGE, H., HAZEN, W., ROSSI, B., and WILLIAMS, R.—*Ibid.* **74** : 1083 (1948).
- (13) BETHE, H.—*Ibid.* **72** : 172A (1947).
- (14) WILLIAMS, R.—*Ibid.* **74** : 1689 (1948).
- (15) EULER, H.—*Z. Phys.* **116** : 73 (1940).
- (16) WOLFENSTEIN, L.—*Phys. Rev.* **67** : 238 (1945).
- (17) BLATT, J. M.—*Ibid.* **73** : 1252A (1948).
- (18) PRESCOTT, J. R.—*Rev. Sci. Instrum.* (in press).
- (19) MONTGOMERY, C. G., and MONTGOMERY, D. D.—*Ibid.* **10** : 350 (1939).
- (20) CORSON, D. R., and BRODE, R. B.—*Phys. Rev.* **53** : 773 (1932).
- (21) SEN GUPTA, R. L.—*Nature* **146** : 65 (1940).
- (22) HAZEN, W. E.—*Phys. Rev.* **67** : 269 (1945).
- (23) LEA, D. E.—*Proc. Camb. Phil. Soc.* **30** : 80 (1934).
- (24) COX, E. F.—*Phys. Rev.* **45** : 503 (1934).
- (25) CARMICHAEL, H.—*Ibid.* **74** : 1667 (1948).
- (26) MONTGOMERY, C. G., and MONTGOMERY, D. D.—*Ibid.* **72** : 131 (1947).
- (27) BAGGE, E.—“Cosmic Radiation.” Ch. XIII. (Dover Publications : New York, 1946.)
- (28) PERKINS, D. H.—*Nature* **160** : 299 (1947).
- (29) ALICHANIAN, A., and ALEXANDRIAN, A.—*J. Phys. U.S.S.R.* **10** : 296 (1946).
- (30) JURITZ, J. W. F., and MOHR, C. B. O.—*Proc. Roy. Soc. A.* **190** : 426 (1948).
- (31) BERNADINI, G., CORTINI, C., MANFREDINI, A.—*Phys. Rev.* **74** : 845 (1948).
- (32) ROSSI, B.—*Rev. Mod. Phys.* **20** : 537 (1948).

# SOLAR RADIO-FREQUENCY RADIATION OF THERMAL ORIGIN

By J. L. PAWSEY\* and D. E. YABSLEY\*

[*Manuscript received January 17, 1949*]

## *Summary*

Observations of solar radiation in the wavelength range from one centimetre to four metres are studied in relation to expected thermal radiation. The data are derived partly from published and partly from unpublished observations.

It is found that a relatively constant component can be identified throughout the whole of this wavelength range despite the complication introduced on the longer wavelengths by the presence of highly variable components. This steady component has the properties expected of thermal radiation and it is concluded that it is, in fact, thermal radiation from the ionized gases of the outer atmosphere of the sun.

The intensity of radiation is found to increase fairly uniformly from that corresponding to black-body radiation at about  $10^4$  °K. at 1.25 cm. to about  $10^6$  °K. at 1.5 m.

The results yield direct confirmation of the hypothesis that the corona has a kinetic temperature of about a million degrees.

## I. INTRODUCTION

The observation by Southworth(1) of radio-frequency radiation from the sun was stimulated by the thought that thermal radiation should extend over the whole spectrum of the infra-red into the radio-frequency range. He found that at wavelengths of 3 and 10 cm. the intensity was equal to that of black-body radiation from a body of the angular diameter of the sun at a temperature of  $18,000$  °K. This temperature is of the same order as the sun's optical temperature of  $6000$  °K. and Southworth considered the observed radiation to be of thermal origin.

The observations which followed at wavelengths between 1 and 5 m.(2, 3) showed that in this range the intensities were much higher than those obtained by extrapolation of the law of black-body radiation from the photosphere. The received intensity was never less than 100 times that corresponding to the optical temperature, and in the presence of certain sunspots large increases in intensity, varying in duration from a few seconds to several days, were observed. The higher intensities correspond to black-body temperatures of up to  $10^8$  or even  $10^{13}$  °K. These are enormously greater than known or suspected temperatures in the sun. This, together with the rapid fluctuations, has been interpreted as indicating the presence of components of non-thermal origin.

This does not preclude the existence of thermal radiation in addition, and studies of the intensity of thermal radiation to be expected from the sun under normal conditions were made independently by Ginsburg(4) and Martyn(5, 6).

\* Division of Radiophysics, C.S.I.R.O.



Each used the Kirchhoff principle relating emissivity and absorptive power, and deduced from the known temperature and electron distribution in the upper solar atmosphere that thermal radiation corresponding in intensity to a temperature of about a million degrees should be expected from the solar corona at wavelengths of a few metres. At shorter wavelengths the corona should become transparent and the intensity fall towards a value corresponding to the temperature of the chromosphere, which is of the order of  $10^4$  °K.

Martyn developed these ideas further than Ginsburg and derived the regions of origin in the solar atmosphere, the distribution of emission across the disk of the sun, and the total intensity at the earth to be expected in the wavelength range from 1 cm. to 30 m. He also discussed the effect of the general magnetic field of the sun on the polarization of this radiation. The intensity previously observed by Southworth at 10 cm., a wavelength where the non-thermal radiation was not evident, agreed with that derived.

Simultaneously with Martyn, Pawsey(7) published evidence concerning the identification of a thermal component at 1.5 m., a wavelength on which the rapidly fluctuating non-thermal component is prominent. He gave a histogram showing the distribution of intensity values on 1.5 m. over a period of six months. It showed that a large proportion fell in a restricted range, in the vicinity of a million degrees, and practically all of the remainder were widely dispersed above this range. This was interpreted as indicating the existence of a relatively constant component together with a highly variable one. Because the constant component was of the intensity predicted by Martyn for thermal radiation it was taken to be thermal.

The mechanism of radiation considered by Martyn in the almost completely ionized gases of the sun's upper atmosphere is, from the classical viewpoint, that associated with acceleration of electrons at electron-ion collisions. From the quantum viewpoint such collisions give rise to "free-free transitions". The emitted radiation has a continuous spectrum but the intensity changes in a peculiar manner with wavelength owing to changes in the relative contributions from the corona, whose electrons are considered to have a kinetic temperature of a million degrees, and the cooler chromosphere which Redman(8) concluded was at a temperature of 30,000 °K.

It is the purpose of this paper to examine the observational evidence now available relating to the thermal component of solar radio-frequency radiation. In Section II, observations of intensity in the metre range of wavelengths are examined and a relatively constant component is isolated from a highly variable one. Section III is concerned with similar observations on shorter wavelengths on which the highly variable component is seldom observed, and Section IV with other observed properties of the more constant component in relation to those expected of thermal radiation. It is concluded that this component is in fact the expected thermal one and its intensity at different wavelengths is given.

Before proceeding it is desirable to mention the method of measurement and of specifying the intensity of solar radiation at a given wavelength. In the observations discussed the quantity measured is the solar power received on a receiver of known bandwidth from a plane-polarized aerial of known effective

area. This leads to a flux density per unit bandwidth of the plane-polarized component. The "noise flux density" of the radiation,  $S$ , is then obtained by multiplying by 2, on the assumption of an equal flux in the component plane polarized at right angles. The intensity may be specified in terms of  $S$  in watts per square metre per cycle per second or, alternatively, in terms of the temperature of a black body subtending the same solid angle as the *visible sun* which would yield the observed radiation intensity. We describe this temperature as the "apparent temperature",  $T_a$ .  $S$  and  $T_a$  are related by the equation

$$S = \frac{2kT_a}{\lambda^2} \Omega_s,$$

where  $k$  = Boltzmann's constant ( $1.38 \times 10^{-23}$  joule/degree),

$\Omega_s$  = the solid angle subtended by the sun (mean  $6.8 \times 10^{-5}$  steradians),

$\lambda$  = wavelength (metres),

$T_a$  is in degrees Kelvin.

## II. SOLAR RADIATION INTENSITIES AT METRE WAVELENGTHS

The series of observations studied are shown in Table 1. The observations consist of daily values of intensity, based on a mean over the hours of observation but neglecting short-term increases. These values are subjected to a statistical analysis which is an extension of that previously used.

The sunset observations of series 1 are plotted against time in Figure 1 (A). It will be noted that there appears to be a base level to which the intensity frequently falls. This is shown better in Figure 1 (B) where the observations have been used to construct a histogram which is seen to be markedly skew, having a sharp cut-off on the low side. It is common for distributions of physical

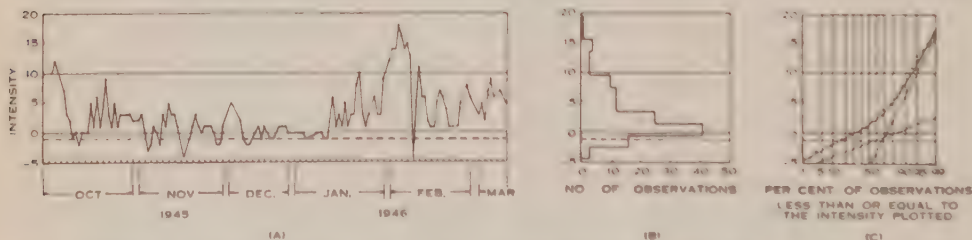


Fig. 1. — Intensity observations of series 1 (a), October 1945–March 1946, 1.5 m. The intensity units are db. above  $10^{-21}$  watts  $\text{m}^{-2}$   $(\text{c/s.})^{-1}$ .

(A) Plot of daily intensities.

(B) Histogram of daily intensities.

(C) Ogive of daily intensities.

quantities to show such a skew distribution when plotted on a linear scale but to be more nearly normal when plotted on a logarithmic scale. The scale of the histogram of Figure 1 (B), being a linear plot of decibels, is logarithmic. Therefore the sharp cut-off at the low intensity end suggests that there is one component of the radiation which is relatively constant while there is another which is highly variable and frequently falls to a small fraction of the first.

TABLE 1  
METRE WAVELENGTH OBSERVATIONS ANALYSED

Series	Observers	Wavelength (m.)	Period of Observations	Place	Hours of Observation	Accuracy (db.)
1 (a)	McCready, Pawsey, and Payne- Scott(3) .. .. .	1.5	Oct. 1945-Mar. 1946	Sydney	Half hour at sunset; most days	$\pm 2$ or 3
1 (b)	McCready, Pawsey, and Payne- Scott(3) .. .. .	1.5	Oct. 1945-Mar. 1946	Sydney	Half hour at sunrise; most days	$\pm 2$ or 3
2	Present authors .. .. .	1.5	Mar.-May 1947	Sydney	Several hours during day; most days	$\pm 1$
3	Lehany and Yabsley(13) .. .. .	1.5	Aug.-Nov. 1947	Sydney	Several hours during day; most days	$\pm 1$
4	Ryle and Vonberg(10) .. .. .	1.71	Dec. 1946-Apr. 1947	Cambridge	Several hours about noon; most days	Not known
5	Ryle and Vonberg(10) .. .. .	3.75	Dec. 1946-Apr. 1947	Cambridge	Several hours about noon; most days	Not known
6	Reber(11, 12, and personal com- munication) .. .. .	1.87	1943-1944	Illinois	Not published. Ref. (11) gives 6 days Nov. 1943- Jan. 1944	Not known

If the less variable component were constant, if there were no observational errors, and if the more variable component fell not infrequently to a negligible value, the cut-off in the histogram would be abrupt at the value of the constant component. It is obvious that these conditions may not be accurately fulfilled, so that the mean value of the relatively constant component is expected to be a little above the lowest observed values as indicated by the dashed lines of Figures 1 (A) and 1 (B).

The best value for this quantity is difficult to estimate, and in an attempt to improve its accuracy the observations of Figures 1 (A) and 1 (B) were replotted on normal probability paper as shown in Figure 1 (C). This paper is ruled in such a way that a normal probability distribution appears as a straight line. The ordinates are intensities (in db.) and the abscissæ the percentage of the total observations included up to the given intensity. Such a curve is called an ogive (see, for example, Worthing and Geffner 9).

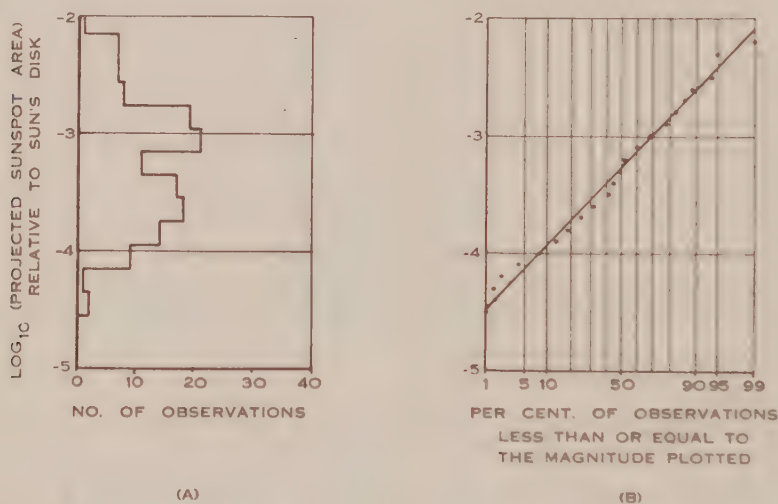


Fig. 2.—Statistics of daily values of sunspot area over the same period as series 1.

(A) Histogram of  $\log$  (sunspot area); steps of 0.2 in  $\log_{10}$ .

(B) Ogive of  $\log$  (sunspot area); steps of 0.2 in  $\log_{10}$ .

The resulting curve of Figure 1 (C) appears to consist not of one but of two intersecting straight lines. The abnormal distribution shown by the histogram and ogive could conceivably be due to a similar abnormality of distribution of solar activity. In an attempt to check such an hypothesis the values for the same period of sunspot area, a quantity known to be closely related to solar noise increases, were plotted in histogram and ogive form as shown in Figures 2 (A) and 2 (B). The scale of sunspot area is logarithmic. The histogram is not skew and the ogive approximates to a single straight line. The abnormality in the distribution of solar noise intensity over this period is therefore not related to a similar abnormality in that of sunspot area.



The shape of the curve of Figure 1 (C) suggests again that the intensity is the sum of two independently variable components, one having an approximately normal distribution on a logarithmic scale but a wide dispersion, the other being either constant or of small dispersion. Theoretical curves of a quantity so compounded were computed numerically (see Figs. 3 (A) and 3 (B)) and conform in shape with the experimental curve. On the basis of this similarity, and taking into account Figures 1 (A) and 1 (B), lines representing the postulated components were drawn in Figure 1 (C). From these lines median values and degrees of dispersion were read off for the more constant component. The degree of dispersion is specified as the range which includes 50 per cent. of the observations. In these and subsequent figures the full line represents the resultant, the dot-dash lines the individual components, and the horizontal dashed line the median of the more constant component. The medians and variabilities of this component have been collected in Table 2, which summarizes the results of analysis of this series and the other observations to be discussed below.

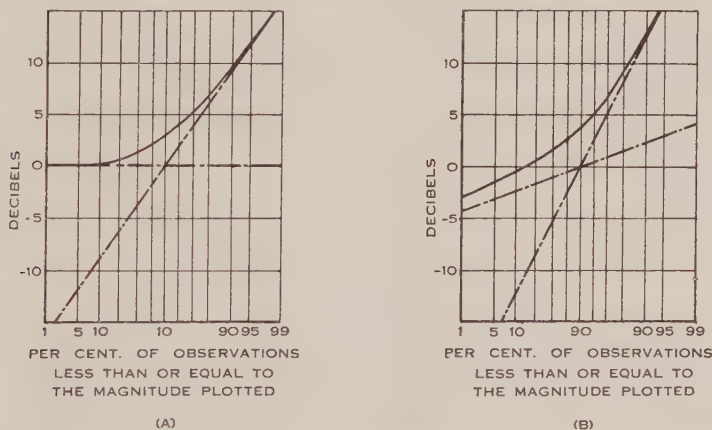


Fig. 3.—Examples of theoretical distributions of a quantity which is the sum of two variable components of widely differing dispersions but each independent and having a normal distribution of its logarithm. The curves are ogives of logarithms of magnitude on normal probability paper.

- (A) Component 1. Magnitude unity ( $=0$  db.), constant.  
 Component 2. Median unity ( $=0$  db.), quartiles at  $\pm 5$  db.  
 (B) Component 1. Median unity ( $=0$  db.), quartiles at  $\pm 1.3$  db.  
 Component 2. Median unity ( $=0$  db.), quartiles at  $\pm 6.7$  db.

The justification for the statistical method used is that the arbitrary assumptions made as to distribution of the postulated components lead to results which are in accord with the observations. The authors do not claim the method is mathematically rigorous but believe that, when used in conjunction with histograms and simple intensity-time graphs as in this paper, it reduces inconsistencies in interpretation.

The sunrise observations of series 1 (B) were similarly treated and the results are included in Table 2. The figure shows similar features and has not been reproduced.

TABLE 2

INTENSITY OF RELATIVELY CONSTANT COMPONENT AT METRE WAVELENGTHS

Series	Observers	Wavelength (m.)	Relatively Constant Component		
			Median Apparent Temp. (°K.)	Variability Range to Include 50% of Observations (db.)	
1 (a)	McCready, Pawsey, and Payne-Scott(3)	1.5	1.0	} 1.2 × 10 <sup>6</sup>	± 1.2
1 (b)		1.5	1.4		± 1.2
2	Present authors .. .. .	1.5	0.7 × 10 <sup>6</sup>		± 0.6
3	Lehany and Yabsley(13) .. ..	1.5	1.0 × 10 <sup>6</sup>		± 0.3
4	Ryle and Vonberg(10) .. ..	1.71	0.6 × 10 <sup>6</sup>		± 0.8
5	Ryle and Vonberg(10) .. ..	3.75	1.3 × 10 <sup>6</sup>		± 0.5
6*	Reber(11, 12, and personal com- munication) .. .. .	1.87	1.8 × 10 <sup>6</sup>		—*

\* Fluctuations in intensity were not observed at this less active period (1943-44) of the solar cycle.

Figures 4 and 5 show the observations of series 2 and 3, taken during two periods of 1947. A larger aerial than that of the earlier measurements was used and the accuracy with which the received intensity could be measured when the sun was "quiet" was greatly increased. This is reflected in a sharper

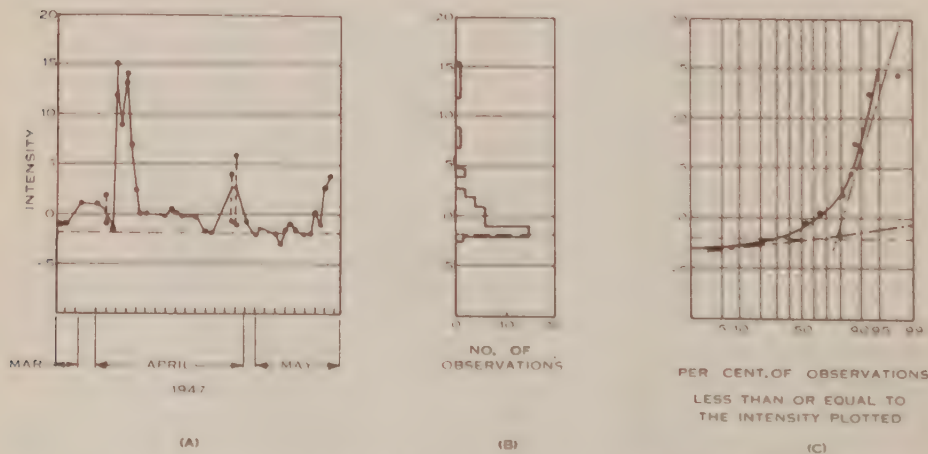


Fig. 4. Intensity observations of series 2, 1.5 m., March-May 1947. The intensity units are db. above  $10^{-21}$  watts  $m^{-2}$   $(c/s.)^{-1}$ .

cut-off at the low end of the intensity scale as illustrated in the histograms of Figures 4 (B) and 5 (B), despite the fact that smaller steps were chosen when constructing these histograms owing to the greater accuracy of the observations. Thus it seems probable that the scatter at the low-intensity end is largely due to experimental errors.

Ryle and Vonberg's(10) observations were submitted to the same statistical analysis, and the results are given in Figures 6 and 7. The distribution of the intensity values is similar to that of the observations made by the present authors and their colleagues.

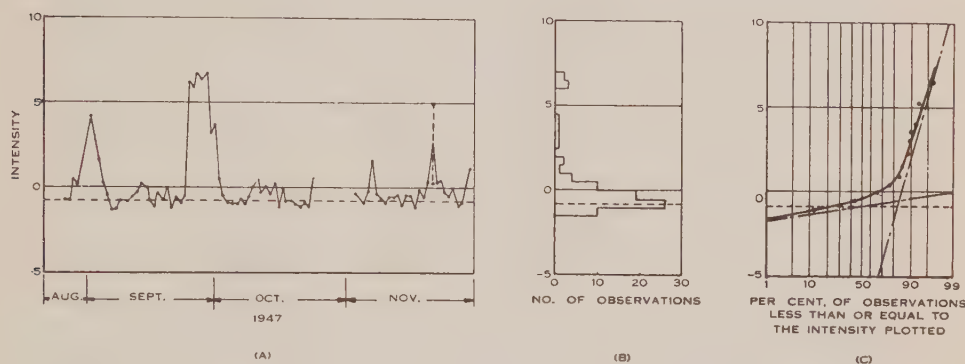


Fig. 5.—Intensity observations of series 3, 1.5 m., August–November 1947. The intensity units are db. above  $10^{-21}$  watts  $\text{m}^{-2}$   $(\text{c/s.})^{-1}$ .

To summarize, all the observations considered above are shown to be consistent with the hypothesis that solar radio-frequency radiation in the metre range of wavelengths consists of two components, one of which is almost constant in intensity over periods of months.

In addition to these observations taken at the present period of great solar activity, observations of solar noise on 160 Mc/s. were made by Reber during the period of low solar activity about the end of 1943. He did not in the original publication of these results(11) reduce the readings to apparent temperatures but has since done so(12), finding an apparent temperature of  $1.8 \times 10^6$  °K. Reber (personal communication) observed no substantial fluctuations and it is reasonable to assume that his measurement relates to the steady component.

### III. SOLAR RADIATION INTENSITIES AT CENTIMETRE AND DECIMETRE WAVELENGTHS

In this wavelength range several extended series of observations are available, as shown in Table 3. The rapid variations characteristic of metre wavelengths occur very rarely and, if such rare occurrences are excluded, all the observations over a period of months fall in a range of intensities of about two to one. This is illustrated in the histograms of Figure 8. The radiation is therefore similar in intensity distribution to the more constant component at metre wavelengths.

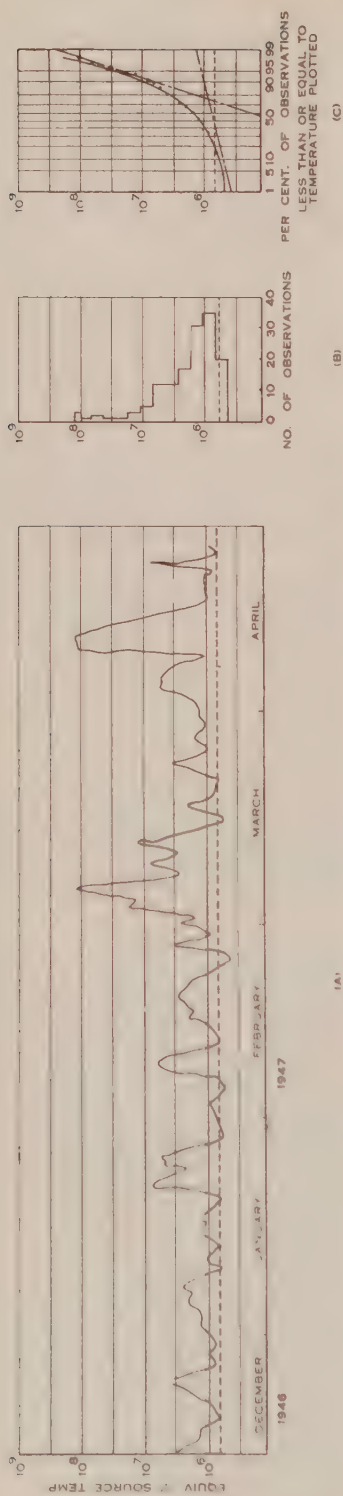


Fig. 6. Intensity observations of series 4 (Ryle and Vonberg) 1.71 m., December 1946-April 1947. The intensities are expressed in terms of the equivalent temperature of a black-body source subtending on angle of half a degree, which differs inappreciably from the "apparent temperature" used here.

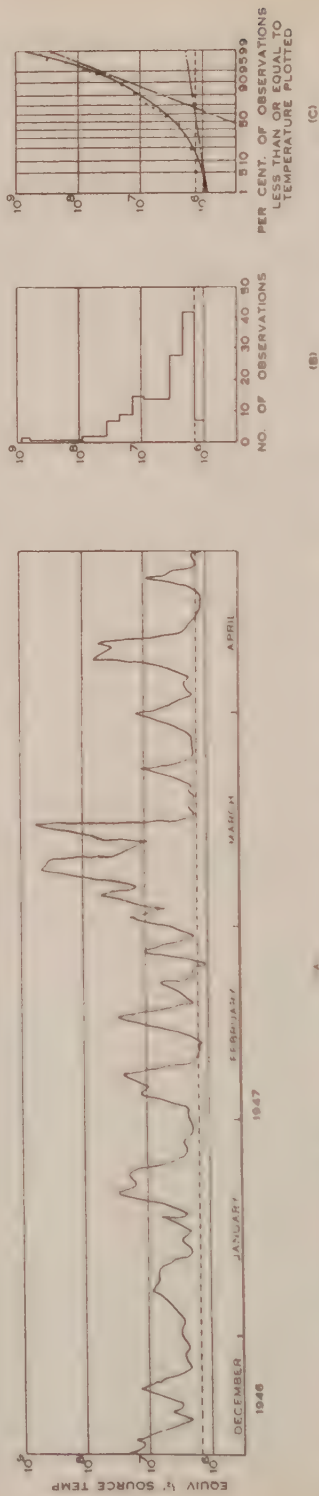


Fig. 7. Intensity observations of series 5 (Kyle and Vonberg) 3.75 m., December 1946-April 1947. The intensities are expressed as in Figure 6.



TABLE 3  
SERIES OF SHORTER WAVELENGTH OBSERVATIONS ANALYSED

Series	Observers	Wavelength (cm.)	Period of Observation	Place	Hours of Observation	Accuracy (db.)
7	Lehany and Yabsley(13)	50	Aug.-Nov. 1947	Sydney	Several hours during day; most days	$\pm 1$
8	Lehany and Yabsley(13)	25	Aug.-Nov. 1947	Sydney	Several hours during day; most days	$\pm 1$
9	Covington(14)	10.7	Feb.-Oct. 1947	Ottawa	Few hours near noon; most days	Not known
10*	Reber(15)	62.5	Aug.-Oct. 1946	Illinois	Not known	Not known
11*	Southworth(1)	10	4 days Aug. 1943	New Jersey	Not known	Not known
12*	Southworth(1)	3	Northern summer 1942 and 1943	New Jersey	Not known	Not known
13*	Piddington and Minnett (personal communication)	1.25	Apr.-Aug. 1948	Sydney	Hour or two; most days	$\pm 0.5$

\* Daily values not available.

Owing to the absence of the intense fluctuating component, it is possible to study details which if they existed at metre wavelengths were completely masked by the more variable component of the radiation. In the observations of series 7, 8, and 9, for which detailed results are available, the observers found small but real variations in intensity which were closely correlated with sunspot

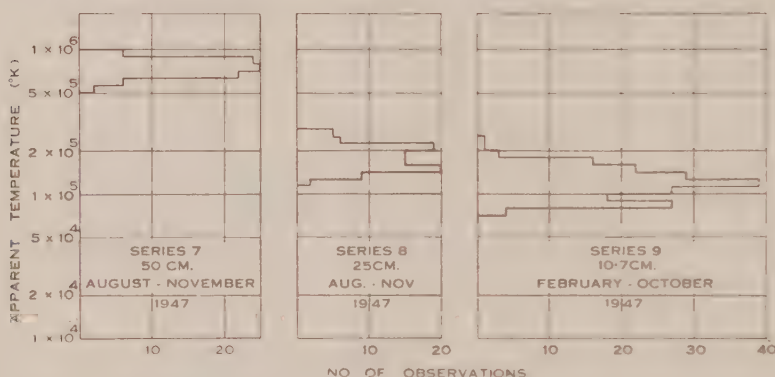


Fig. 8.—Histograms of intensity observations of series 7, 8, and 9.

activity. Figure 9 illustrates the correlation with projected sunspot area for each case. Dashed lines have been drawn indicating the trend of the relation between sunspot area and mean intensity. The correlation diagrams indicate clearly a “background level” in the absence of sunspots added to a contribution which, if the cases of exceptionally great sunspot area are excluded, is approximately proportional to sunspot area. We quote the background level rather

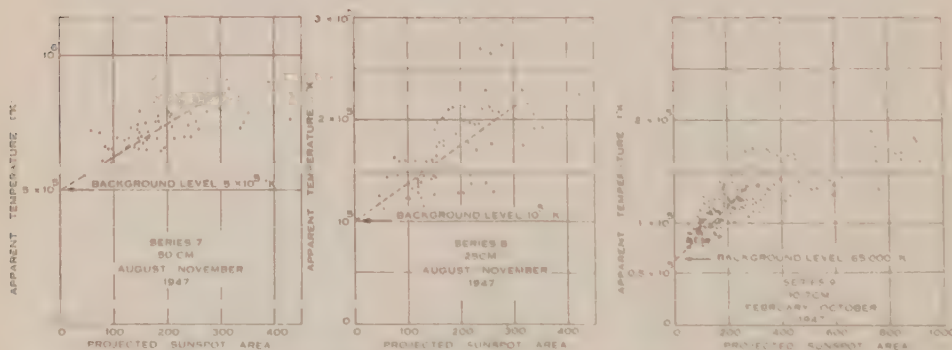


Fig. 9. Correlation diagrams relating intensity and sunspot area for series 7, 8, and 9. Sunspot area units are  $10^{-5}$  of solar disk.

than the mean level of the radiation since it should relate more closely to the computations made for the “quiet” sun. In addition, it is not yet known if the increases are of thermal origin. Lehany and Yabsley(13) give reasons for supposing that on decimetre wavelengths the increases of radiation, which are correlated with sunspot area, are distinct from the larger ones observed on

metre wavelengths, and are due to increased thermal emissivity of the corona caused by sunspot magnetic fields. They indicate, however, that an alternative mechanism remains a possibility. The background levels deduced are given in Table 4.

Similar correlation diagrams were drawn for metre-wavelength observations of the previous section but the relationship between sunspot area and solar noise intensity was found to be insufficiently close to justify similar deductions concerning the background level. These diagrams are not reproduced.

The observers of series 10 to 13 found no evidence for appreciable fluctuations and their measured intensities are given in Table 4 as a measure of the background level.

TABLE 4  
"BACKGROUND LEVEL" OF INTENSITY AT SHORTER WAVELENGTHS

Series	Observers	Wavelength (cm.)	Apparent Temp. of Background Level (°K.)	Derivation
7	Lehany and Yabsley(13) ..	50	$0.5 \times 10^6 \pm 20\%$	Correlation diagram
8	Lehany and Yabsley(13) ..	25	$1.0 \times 10^6 \pm 20\%$	
9	Covington(14) .. ..	10.7	$6.5 \times 10^4$	
10	Reber(15) .. .. .	62.5	$1.0 \times 10^6$	Mean of a series with trivial variations
11	Southworth(1) .. ..	10	$1.8 \times 10^4$	
12	Southworth(1) .. ..	3	$1.8 \times 10^4$	
13	Piddington and Minnett (personal communication)	1.25	$1.0 \times 10^4 \pm 10\%$	
14	Covington(18) .. ..	10.7	$5.6 \times 10^4$	Eclipse observation
15	Sander(17) .. .. .	3.2	$2.2 \times 10^4$	
16	Dicke and Beringer(16) ..	1.25	$1.0 \times 10^4$	

In addition to the seven series of observations already quoted, observations taken on different wavelengths during eclipses can supply information concerning the base level of intensity provided supplementary evidence can be given to show that the intensity measured was not a temporary high value due to sunspot activity. On 1.25 cm. Dicke and Beringer's(16) 1945 observation of an apparent temperature of  $10^4$  °K. was shown to be associated with a reasonably uniform distribution over the solar disk so that it would appear representative of the "quiet" sun. This result is supported by unpublished measurements on this wavelength by Piddington and Minnett over the period April-August 1948. They found no fluctuations greater than  $\pm 10$  per cent. and their mean apparent temperature was also  $10^4$  °K. with an error believed less than 10 per cent.

On 3.2 cm. Sander's(17) 1945 observation of  $2.2 \times 10^4$  °K. is again probably representative since the radiation was shown to originate fairly uniformly over the sun's disk.

On 10.7 cm. Covington's(18) 1946 measurement showed an intense source over a sunspot group. This was subtracted from the total intensity to give 56,000 °K. for the average for the remainder of the disk, over which the distribution was reasonably uniform. These three observations are added to Table 4 with identifying serial numbers.

#### IV. FURTHER CHARACTERISTICS OF THERMAL RADIATION

In this section further observable characteristics of solar radiation are discussed in relation to those expected of thermal radiation.

Firstly, thermal radiation should show the random waveform characteristic of fluctuation noise. Only qualitative observations obtained aurally on loud-speakers or visually on cathode-ray oscillographs are available. They indicate that all "solar noise" has this property.

Secondly, it was predicted by Martyn that thermal radiation from the whole of the quiet sun would show no circular polarization. He considered that at certain wavelengths the general solar magnetic field would lead to the radiation from a single hemisphere having a degree of circular polarization, but that the radiation from the other hemisphere would be of opposite sense of rotation so that the resultant of the two should have none.

It is well known that "solar noise" on metre wavelengths is frequently circularly polarized but these observations refer to high-intensity radiation. Observations taken in this Laboratory on 1.5 m. have consistently shown random polarization at times of low intensity when the thermal component is expected to be dominant. Observations on shorter wavelengths are not available.

Thirdly, thermal radiation is expected to originate more or less uniformly over the sun's disk. Detailed distributions predicted by Martyn include a bright ring around the rim in the decimetre range of wavelengths. Observations of the time-variation of intensity at eclipses of the sun have been made by Dicke and Beringer at 1.25 cm., by Sander at 3.2 cm., and by Covington at 10.7 cm., as previously mentioned. The observations consistently indicated that at these short wavelengths the observed radiation originated fairly uniformly over the solar disk; Covington also observed noticeable enhancement over a sunspot group. There may have been an enhancement at the limb but the evidence was not conclusive. On metre wavelengths estimates of position and size of the radiating areas on the sun have been made using interference methods. Intense radiation has frequently been found to originate in small areas on the sun's disk, but observations in this Laboratory have consistently shown that when the intensity is low, so that the thermal component might be expected to be dominant, the source is found to approximate to the whole solar disk. No precision was possible in these particular observations because of the frequent presence of non-thermal radiation during the period of the observations.

In all these respects, therefore, the observations are consistent with the hypothesis that in the metre range of wavelengths at times of low intensity, and at shorter wavelengths at nearly all times, solar radiation is mainly thermal in origin.



## V. DISCUSSION OF RESULTS AND CONCLUSIONS

Observations of solar radiation intensities throughout the whole of the wavelength range from 1 cm. to 4 m. give clear evidence for a relatively constant component at each wavelength. There are also more variable components with which this paper is not concerned.

The magnitude of the relatively constant component at different wavelengths, as derived from measurements by various workers, is shown in Figure 10. The points are identified by the series numbers of Tables 1-4 and by the year of observation. This graph contains three classes of points: open circles, which

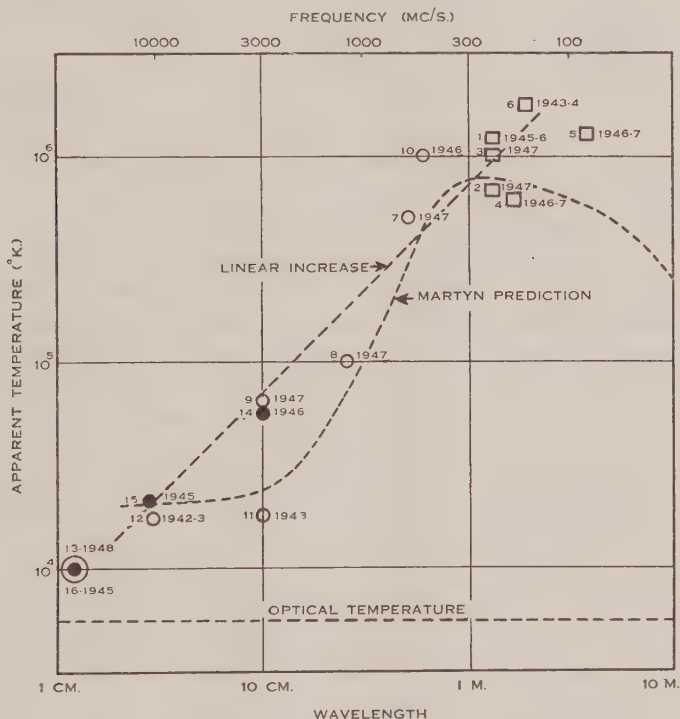


Fig. 10.—Summary of intensities of relatively constant component at different wavelengths. Points are identified by the series numbers of Tables 1-4, and by the year of observation.

○ denotes "background level" derived from series of observations.

● denotes "background level" derived from eclipse observations.

□ denotes "median" value of metre wavelength observations.

are "background level" determinations in the decimetre range, as explained in Section III; black disks, which are estimates of the same quantity derived from eclipse observations, as explained in Section III; and squares, which are estimates of the median value of the more constant component in the metre range of wavelengths, as described in Section II.

The inconsistency introduced by quoting first a "background level" and then a "median" is necessary because in the metre range of wavelengths the

non-thermal component prevents analysis of the data with the object of determining a possible corresponding background level. However, the dispersions given in Table 2 indicate that this inconsistency should not introduce errors of more than about 20 per cent., which is not serious.

Figure 10 also shows Martyn's predicted apparent temperatures. It will be seen that over a range of intensities of 100/1 there is substantial agreement between the observations and the predictions. Further, it was shown in Section III that available observations of waveform, polarization, and region of origin on the sun of the relatively constant component were consistent with a thermal origin. There is, therefore, no doubt that the radiation studied here is mainly thermal radiation from the upper solar atmosphere.

The observed points at certain wavelengths show a considerable scatter. Thus at 10 cm. Southworth's observation in 1943 gave 18,000 K. while Covington's in 1946 gave 56,000 K., and in 1947, 65,000 K. The discrepancy of between 3 and 4 to 1 between Southworth and Covington may be due to observational errors or to long-term changes in solar emission. It will be noted that 1943 was a period of low solar activity and 1947 had reached the peak of an exceptionally active sunspot cycle. Current measurements by Piddington (personal communication, November 1948) are in general agreement with Covington's values, so that it appears that the latter are substantially correct. However, it is impossible now to check Southworth's measurements and the decision as to whether the level of thermal radiation on 10 cm. falls at times of low solar activity must await measurements at a later stage in the sunspot cycle. It should be noted that the radiation under discussion is not that directly associated with sunspots since the procedures adopted in Section III eliminate this. An analogous scatter of observations occurs in the vicinity of 1.5 m. but in this case the observation at a period of low sunspot activity, which cannot be checked, is a high value.

Taking the more recent values, it is of interest that the observations are moderately well represented by a simple linear increase of apparent temperature with wavelength from  $10^4$  K. at 1.5 cm. to  $10^6$  K. at 1.5 m., as indicated by the dashed sloping line of Figure 10.

The outstanding conclusion concerning the nature of the solar atmosphere which may be drawn from these observations is that they offer direct confirmation of the existence of kinetic temperatures of the order of a million degrees in the corona.

A corresponding deduction concerning the supposed excess temperature of the chromosphere over that of the photosphere requires more careful consideration than can be given here. It was pointed out by Martyn that for waves of decimetre wavelength and lower the corona would be nearly transparent, so that in this range radiation from the chromosphere should reach the earth without appreciable attenuation. Following this thought, the observation that the observed radio intensities are consistently above those appropriate to 6000 °K. would appear to indicate that the temperature of the chromosphere is similarly above it. Martyn, however, in his approximate treatment neglected to take account of the emission from the corona at such a wavelength. Because

it is nearly transparent the emissivity is likewise small, but it is not zero, and the very great temperature could cause appreciable emission which would add to that from the chromosphere. Quantitative discussion of this effect is necessary to reach a conclusion.

On a wavelength of 1.25 cm. the observed intensity, corresponding to  $10^4$  °K., falls below the temperature of 30,000 °K. derived by Redman for the chromosphere. The 1.25 cm. intensity is one of the best established observations and is evidence that waves on this wavelength reach us from levels in the solar atmosphere which are at a temperature of not more than  $10^4$  °K.

On the long wavelength side Martyn's curve shows a decrease due to a predicted decrease in emissivity of the corona. The observations do not give evidence of such a decrease and measurements at longer wavelengths would be of great interest.

## VI. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Radiophysics, C.S.I.R.O. The authors wish to acknowledge the encouragement and help given by Dr. E. G. Bowen, Chief of the Division, and to thank Dr. J. H. Piddington and H. C. Minnett for making available their unpublished data. Sunspot data were provided by the Commonwealth Observatory through the courtesy of Dr. R. v. d. R. Woolley, Commonwealth Astronomer, and Dr. C. W. Allen.

## VII. REFERENCES

- (1) SOUTHWORTH, G. C.—*J. Franklin Inst.* **239** : 285-97 (1945); and correction. *Ibid.* **241** : (1946).
- (2) APPLETON, E. V., and HEY, J. S.—*Phil. Mag.* **37** : 73-84 (1946).
- (3) MCCREADY, L. L., PAWSEY, J. L., and PAYNE-SCOTT, RUBY.—*Proc. Roy. Soc. A* **190** : 357-75 (1947).
- (4) GINSEBURG, V. L.—*C.R. Acad. Sci. U.R.S.S.* **52** : 487-90 (1946).
- (5) MARTYN, D. F.—*Nature* **158** : 632-3 (1946).
- (6) MARTYN, D. F.—*Proc. Roy. Soc. A* **193** : 44-59 (1948).
- (7) PAWSEY, J. L.—*Nature* **158** : 633-4 (1946).
- (8) REDMAN, R. O.—*Mon. Not. Roy. Astr. Soc.* **102** : 140 (1942).
- (9) WORTHING, A. G., and GEFFNER, J.—"Treatment of Experimental Data." pp. 179-82. (John Wiley : New York, 1943.)
- (10) RYLE, M., and VONBERG, D. D.—*Nature* **160** : 157-9 (1947).
- (11) REBER, G.—*Astrophys. J.* **100** : 279-87 (1944).
- (12) REBER, G., and GREENSTEIN, J. L.—*Observatory* **67** : 15-26 (1947).
- (13) LEHANY, F. J., and YABSLEY, D. E.—*Aust. J. Sci. Res. A* **2** : 48-62 (1949).
- (14) COVINGTON, A. E.—*Proc. Instn. Radio Engrs.* **36** : 454-7 (1948).
- (15) REBER, G.—*Nature* **158** : 945 (1946).
- (16) DICKE, R. H., and BERINGER, R.—*Astrophys. J.* **103** : 375-6 (1946).
- (17) SANDER, K. F.—*Nature* **159** : 506 (1947).
- (18) COVINGTON, A. E.—*Ibid.* **159** : 405-6 (1947).



# BURSTS OF SOLAR RADIATION AT METRE WAVELENGTHS

By RUBY PAYNE-SCOTT\*

[Manuscript received January 5, 1949]

## Summary

The variable components of solar radiation observed on frequencies of 85, 65, 60, and 19 Mc/s. over a period of nine months have been studied. It is shown that two types of variable high-intensity radiation can be distinguished. One of these, referred to as the "enhanced level", is circularly polarized. The other, a particular type of short duration increase, is not circularly polarized and such an increase is called an "unpolarized burst".

These bursts tend to occur nearly simultaneously over a range of frequencies. They decay exponentially and double-humped bursts are common. Their characteristics are shown to conform broadly with the hypothesis that the bursts originate in localized transitory disturbances in the high corona which radiate over a wide frequency range. The decay constant is correctly predicted on the assumption that it is the decay constant of the excited medium in the region of origin, and the double peaks on the assumption that the second peak is an "echo" of the disturbance after reflection at the appropriate lower level in the corona.

The occurrence of time delays between the arrival of "corresponding" unpolarized bursts on different frequencies is confirmed, the higher frequency commonly arriving earlier, with delays of about 0.7 second between 85 and 60 Mc/s. and 9 seconds between 60 and 19 Mc/s. There is a good correlation between major radio "fade-outs" and large bursts on these frequencies.

## I. INTRODUCTION

The most striking feature of the radiation now known to be emitted by the sun on metre wavelengths is its great variability both over long periods and over a few seconds(1, 2). Normally the intensity is low, corresponding to black-body radiation from the solar disk at a temperature of about a million degrees(3). This low intensity is presumed to be due to thermal radiation. However, at times much higher intensities are produced. Radiation of up to a thousand times this base intensity has been shown to originate in the neighbourhood of large sunspot groups(2), reaching its maximum during meridian transit of the spot(4). Short-term increases in intensity have also been observed(2), lasting for seconds or minutes and exceeding the thermal intensity by up to a million times. These increases are not of terrestrial origin(4) but are believed to arise in the sun. The largest are sometimes associated with flares(4, 5, 6, 7). It has been observed that the rapid variations in intensity are often very different on different frequencies(8, 9), and that when they are similar they tend to occur not quite simultaneously but in the sequence high frequencies before low fre-

\* Division of Radiophysics, C.S.I.R.O.



quencies(9). Several observers(10, 11, 12) have recorded circular polarization of the high-intensity radiation.

For the past nine months we have been recording solar noise continuously on 85 and 60 Mc/s., with some observations on 65 Mc/s. Receiving equipment on 19 Mc/s. designed for moon-echo work has been maintained at the same station by F. J. Kerr and his colleagues of this Laboratory, and he has kindly provided a number of records of solar radiation on this frequency. On the basis of our observations over this period, we have demonstrated the existence of two distinct types of variable high-intensity radiation. These are termed "enhanced radiation" and "unpolarized bursts" respectively. The present paper gives the characteristics of the two types, and then discusses in more detail the characteristics of the "unpolarized burst" type, particularly the shape of individual bursts, their relative intensities and time of appearance on different frequencies, and their relation to radio fade-outs. The information obtained on day-to-day variations will be given in a separate paper.

## II. EQUIPMENT AND RECORDING TECHNIQUES

The observing period was from January to September 1948. The aerials used on 85, 65, and 60 Mc/s. were single horizontally-polarized Yagis so mounted that they could be directed at the sun. Circular polarization was studied only on 85 Mc/s., a pair of crossed Yagis being used. By inserting a movable quarter-wave of cable in the feed to one Yagi or the other, either the right-handed or the left-handed circularly-polarized component of the incident radiation was received. The aerial used on 18.3 Mc/s. was a plane-polarized broadside array, whose direction was normally vertical but could be changed slightly by phase switching. On 19.8 Mc/s. a fixed rhombic with a bearing  $60^\circ$  E. of N. and elevation  $15^\circ$  was used. Both of these frequencies are referred to as 19 Mc/s. in this paper. Since neither of the 19 Mc/s. aerials could be directed exactly at the sun, intensity measurements made on them were subject to correction factors.

The receivers on the higher frequencies were standard radar receivers with a bandwidth of about 1.5 Mc/s. On 19 Mc/s. a communication receiver was used, with a bandwidth of 5 kc/s.

Usually the receiver output was recorded on an Esterline-Angus recording milliammeter. Time marks were put on all records simultaneously through relays actuated by a common switch. The switch operated from a synchronous motor running from the 50-cycle mains, and could be adjusted to operate at various time intervals down to half a minute; it could also be operated manually. The time-constant of the recorders is about 0.2 second, which sets a limit to the detail that can be observed. At fast speeds (12 in./min. and 6 in./min.) the reading accuracy of the recorder is about  $\pm 0.2$  second, at  $\frac{3}{4}$  in./min. it decreases to about  $\pm 2$  seconds, and at slower speeds the accuracy is proportionally less.

To check the effect produced by the recorder on the shape of solar noise variations, the video output of the receivers could be fed through an amplifier

to a long-persistence cathode-ray oscillograph\* with a time-base sweep of 6 seconds, the fly-back occupying about one second of this. To avoid too much blurring of the line on the screen, the video response was cut off at about 100 c/s. The screen could be photographed by a ciné-camera geared to the time-base switch, so that each frame corresponded to a single sweep of the time-base. By feeding the outputs of two receivers through a mechanical switch operating at 25 c/s. the variations on two frequencies could be displayed simultaneously on the cathode-ray screen in the form of two traces slightly displaced from one another.

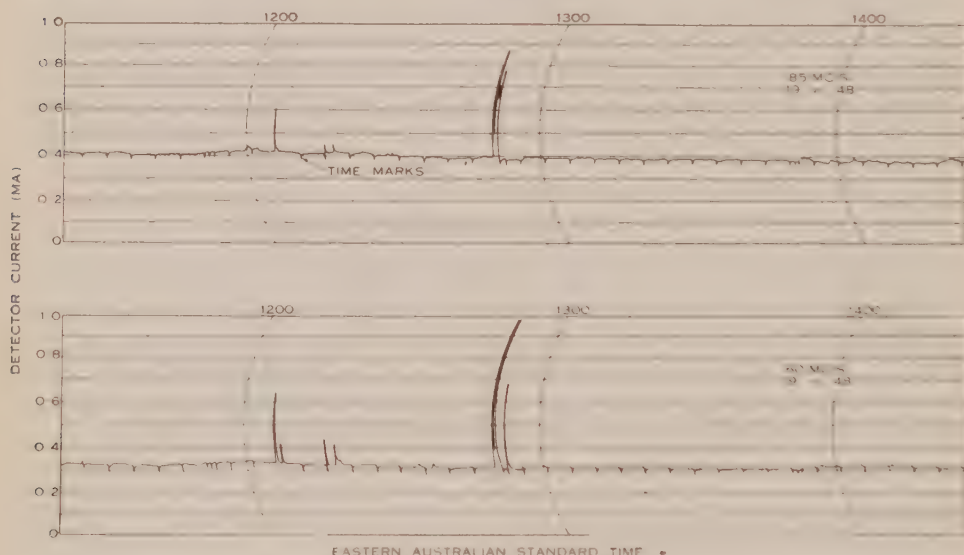


Fig. 1.—Typical recordings of unpolarized bursts on 60 and 85 Mc/s.

### III. THE TWO TYPES OF VARIATION

Figure 1 shows a few hours of a typical day's record on two frequencies. At first, the intensity of solar radiation is almost undetectable at the very low thermal level. Suddenly a burst of radiation occurs, lasts a few seconds, and then disappears. In the next hour several such bursts occur, some reaching a very high intensity. Each lasts only a few seconds and then disappears. In extreme cases, called "outbursts" by Allen(4), the level may remain high for some minutes. Such an outburst is shown in Figure 2. However large the burst or the outburst, its duration is relatively short. Soon the intensity falls to its original level, and may remain there for hours. The burst is a transitory phenomenon.

Much more rarely quite a different phenomenon occurs. The intensity reaches a high level and remains there for hours or days on end: there are continual fluctuations in intensity, both long-term and short-term. The short-term increases are somewhat similar to the bursts described above, but usually

\* This oscillograph and its associated camera were made available by F. J. Kerr.

have a lower ratio of maximum to background level. This type of radiation will be called "enhanced radiation", adopting the term used by Pawsey(13). Superposed on it there may be bursts. Figure 3 shows this type of radiation.

One would be tempted to assume that "enhanced radiation" is merely a continuous succession of the kind of bursts shown in Figure 1 and that there is no essential physical difference between the two types of variation, were it not for one outstanding fact. The enhanced level is circularly polarized. There

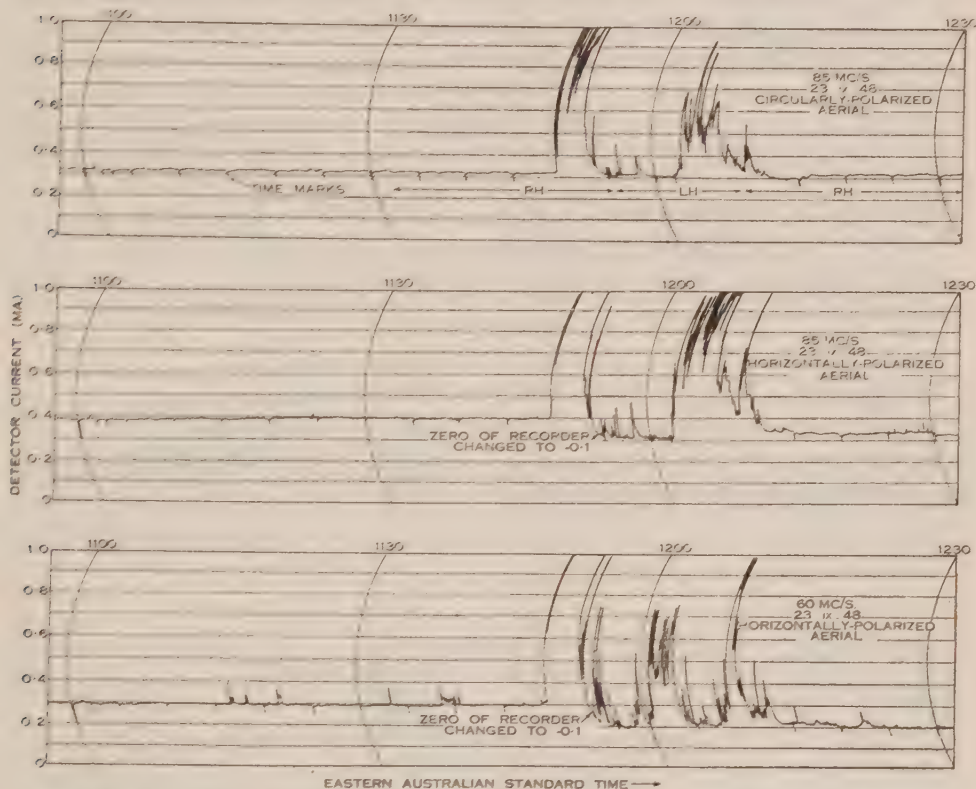


Fig. 2.—Recordings of a typical outburst on 60 and 85 Mc/s. A radio fade-out began at 1151 E.A.S.T.

may be short periods during which the polarization is indefinite, either because two sources of opposite polarization are superposed or because the radiation is linearly or randomly polarized, but for the greater part of its life the enhanced level shows circular polarization of one sense or the other. The type of burst in Figure 1 shows no circular polarization, even when superposed on an enhanced level which is polarized.

The method of ascertaining polarization was to operate simultaneously two 85 Mc/s. receivers, one fed from the horizontally-polarized Yagi and the other from the crossed Yagi, set to receive either the right-handed or the left-handed component, the polarization being changed at intervals of a few minutes. A change in amplitude relative to the horizontally-polarized aerial indicates



circular polarization. Figure 3 shows the type of record obtained when bursts are superposed on an enhanced level. Most of the variations are in the enhanced level, which shows left-handed polarization. An unpolarized burst occurs about 1325. The absence of circular polarization in an outburst is shown by comparing the two 85 Mc/s. records of Figure 2.\* Four other outbursts examined for polarization also showed no signs of circular polarization, so it seems reasonable to assume that outbursts also belong to the class of unpolarized bursts.

The term "burst" has been used to denote short-term increases of any kind. Thus Martyn(10) and Ryle and Vonberg(14) have stated that "bursts" show the same circular polarization as the background level. We interpret these

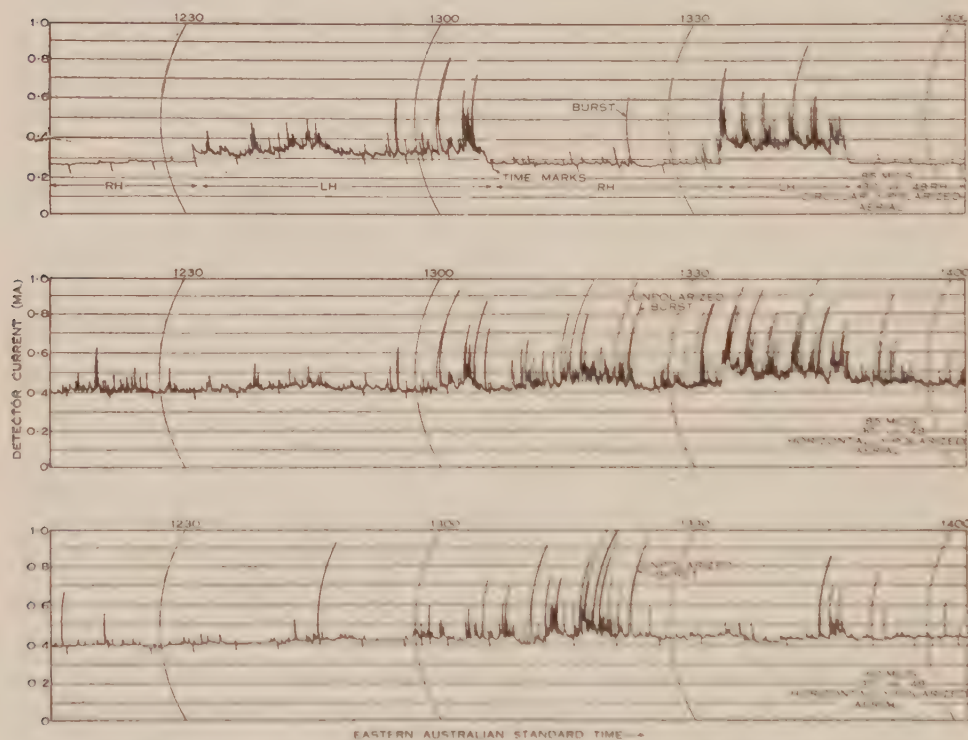


Fig. 3.—Typical recordings of enhanced radiation on 60 and 85 Mc/s.

"bursts" as short-duration variations in the enhanced radiation. Bolton, in unpublished observations quoted by Pawsey(13), has observed that some "bursts" occurring in the absence of enhanced radiation were not circularly polarized. These would appear to be examples of the type considered here. So also is the unpolarized outburst discussed in our letter(9) on time-delays. To avoid confusion we shall henceforth refer to those sudden increases which show no circular polarization as "unpolarized bursts", or "unpolarized outbursts" for those of high intensity and long duration.

\* Note that the sensitivity is greater on the horizontally-polarized system.



This distinction between "unpolarized bursts" and variations in enhanced radiation makes it possible to clarify statements(8, 9) that there is very little correspondence between bursts on different frequencies, only the largest corresponding. Inspection of Figures 1, 2, and 3 shows that such a lack of correspondence is characteristic of the variations in the enhanced radiation but not of the unpolarized bursts, which show a very good correspondence on different frequencies, though their shapes and relative amplitudes may vary considerably. The correspondence between these bursts on different frequencies is further illustrated in Figure 4. It remains true for these bursts that the closer the frequencies and the larger the bursts, the closer their relationship. Corresponding bursts do not appear to skip frequencies. Thus, if a burst appears on 85 and 19 Mc/s., there will be a corresponding burst on 60 Mc/s. However, they may be confined to a particular frequency range, either high or low.

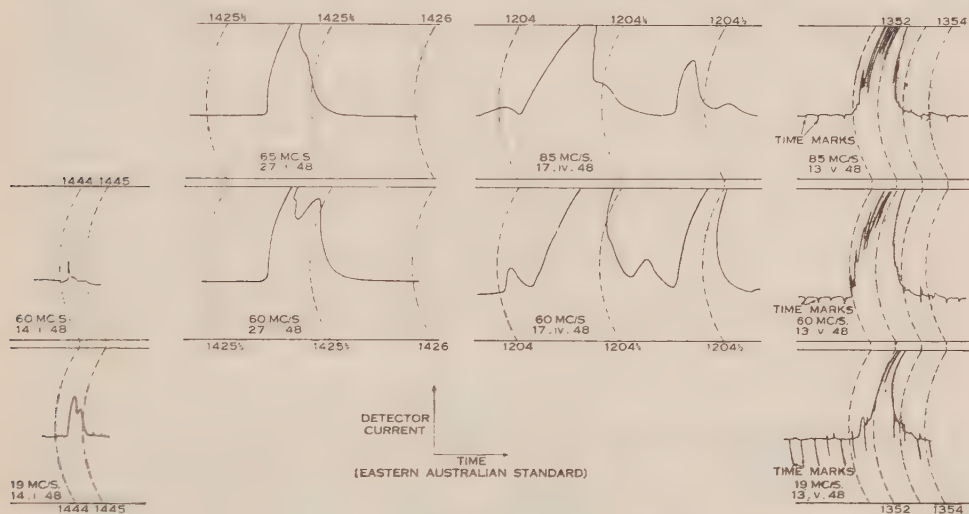


Fig. 4.—Recordings of typical unpolarized bursts.

Although unpolarized bursts seem to be about as common on 19 Mc/s. as on 60 and 85 Mc/s., on only one occasion (August 5–10) did enhanced radiation appear on 19 Mc/s. At other times of enhanced radiation on the other frequencies the 19 Mc/s. records showed only isolated bursts. On the other hand, the intensity of the enhanced radiation on 60 Mc/s. is usually about twice that on 85 Mc/s.

#### IV. THE CHARACTERISTICS OF UNPOLARIZED BURSTS

##### (a) *Relative Intensity on Different Frequencies*

Figure 5 shows a histogram of the relative intensity (ratio of peak energy flux incident on unit area of the earth per unit frequency interval) at 60 and 85 Mc/s. of a number of corresponding unpolarized bursts. The sample does not include very large bursts that drive the recorder off scale, but otherwise is typical. The histogram shows that the lower frequency normally has the higher intensity,

the average ratio of intensity on 60 Mc/s. to that on 85 Mc/s. being about 2 to 1—approximately the ratio of the square of the corresponding wavelengths.

On 19 Mc/s. the interpretation of the corresponding measurements is rendered uncertain by lack of knowledge of attenuation in the ionosphere. Moon-echo measurements on the same equipment by F. J. Kerr indicate that this may be considerable. The bursts measured show an erratic scatter in intensity ratios for 60 and 19 Mc/s., suggesting some variable effect such as ionospheric absorption. On the average, the peak intensity measured at the earth's surface is about equal for corresponding bursts on these two frequencies.

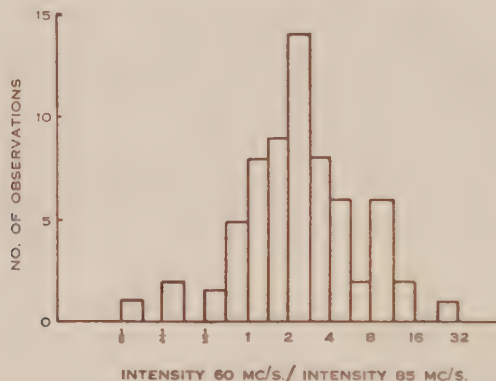


Fig. 5.—Relative intensity of corresponding bursts on 60 and 85 Mc/s.

### (b) *The Shape of Unpolarized Bursts*

A characteristic unpolarized burst shows a finite rise time, rounded top, and slow decay, reminiscent of the transient response of a medium with a natural resonant frequency. The bursts shown in Figure 4 exhibit this characteristic shape. Comparison of about twenty bursts photographed on the cathode-ray screen with the corresponding recorder traces showed that there are no rapid variations in the burst envelope beyond the capacity of the recorder at high speed, so that the recorder produces an adequate representation of the burst envelope.\*

On the assumption that, although the initiating impulse may have a complex form, the rate of decay is likely to be characteristic of the medium rather than of the initiating disturbance, the decay rate of about 100 unpolarized bursts on 85 and 60 Mc/s. recorded at 12 in./min. has been measured. The bursts measured were mainly those used in the time-delay measurements described later. Those chosen all occurred on both frequencies and had a sufficiently long undisturbed decay period without interruption by a subsequent burst.

\* In addition, about four hours of enhanced radiation were photographed. The rate of decay of variations in this radiation is about twice as fast as that of unpolarized bursts, but comparison of Esterline-Angus records with the photographs shows that the recorder produces an adequate record of the envelope of this type of variation also.

The curved coordinates of the recording paper and the characteristics of the linear detector of the receiver were taken into account in determining the decay rates. The base level of intensity was taken as the level before that particular group of bursts appeared. The time-constant of the recorder was neglected, as it is much less than the decay time of the bursts and hence does not seriously affect them. For the same reason, receiver time-constants were neglected.

The decaying portion of the bursts was found to fit very well a curve of the form

$$\left. \begin{array}{l} E = E_0 e^{-\frac{1}{2} \nu t} \\ S = S_0 e^{-\nu t} \\ P = P_0 e^{-\nu t} \end{array} \right\} \dots\dots\dots (1)$$

where  $E$ ,  $S$ , and  $P$  are respectively the electric field in the bursts, the incident power, and the receiver output power due to the burst. An analysis of the measured values of the damping constant  $\nu$  is given in Figure 6 (A). This shows for both frequencies a fairly symmetrical distribution round a median value of about  $0.6 \text{ sec.}^{-1}$

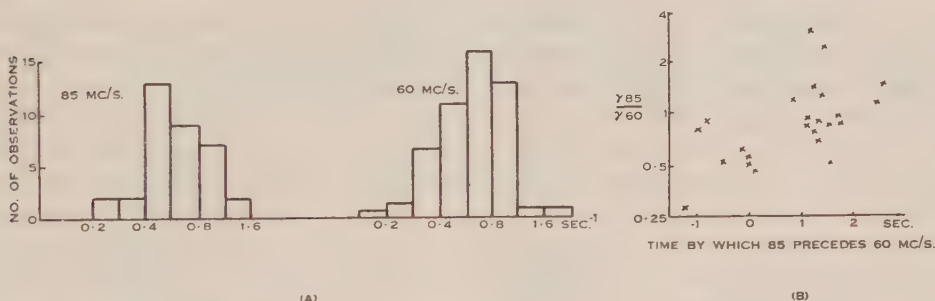


Fig. 6.—(A) Distribution of damping constants.

(B) Scatter diagram of damping constant ratios against time-delays for 85 and 60 Mc/s.

Williams(15) has made similar measurements of decay rates on 75 Mc/s. Converting his half-life values to decay constants gives a most common value for  $\nu$  of  $0.7 \text{ sec.}^{-1}$  with a range from  $0.3$  to  $1.8$ , in good agreement with our results.

A corresponding pair of bursts on the two frequencies normally have about the same damping constant. This can be seen from the scatter diagram of Figure 6 (B). Ignoring for the moment the points on the left-hand side corresponding to bursts for which 60 Mc/s. precedes 85 Mc/s. (see Section IV (c)), the remainder nearly all have a ratio of damping constants between  $0.7$  and  $1.5$ , clustering about the value  $1$ .

There is no marked connexion between the rate of decay and the intensity of the burst, small bursts and outbursts having the same average rates. Occasionally two successive or overlapping bursts have appreciably different decay rates.

The damping constant of bursts on 19 Mc/s. has been measured for only a few samples, and the accuracy of measurement is less, as the records are at a slower speed. However, these bursts have a markedly slower decay rate, their damping constant being of the order  $0.25 \text{ sec.}^{-1}$ .



The interpretation of these damping constants is discussed in Section V.

One further feature of burst shape is significant. The single burst is very uncommon. Much more frequent in occurrence is the double hump apparent in the first two examples of Figure 4. Complex groups of bursts also tend to show double humps in the individual components. Of 45 bursts chosen at random from those measured for time-delays\* (see next Section), 21 were definitely double-humped, and only 5 single. The remaining 19 were too complex and spent too much time off scale for the presence or absence of double humps to be ascertained. Of the 21 double-humped bursts, in 17 the earlier peak and in 4 the second peak was the larger. The ratio of the peak intensity (energy flux) in the two humps had an average value of about 4 for 85 and 60 Mc/s. The time interval between peaks varied from about 2 to 10 seconds. It must be remembered that bursts less than 2 seconds apart would appear as a single burst, so that some of the apparently single bursts may consist of such closely-spaced bursts. Such a hump often appears as a slight change of slope in the decaying portion.

The strong tendency towards double-humped bursts, with the second peak normally the smaller, suggests that the second peak may be an echo of the original disturbance. This possibility is discussed in Section V. On the assumption that the decay rate is a property of the medium, one would then expect double-humped bursts, which would arise outside the assumed echoing layer, to have a slower decay rate than single bursts, which would arise close to this layer. Also one would expect the two parts of a double-humped burst to have the same decay rate. Measurement on 7 single bursts gave decay constants of 0.6 to 1.3 sec.<sup>-1</sup>, with a mean value of 1.0 sec.<sup>-1</sup>. Similar measurements on 14 double-humped bursts gave a range of 0.4 to 1.0, with a mean of 0.7 sec.<sup>-1</sup>. Thus single bursts tend to decay more rapidly than double-humped bursts and also more rapidly than the average of all bursts measured. The comparison of decay rates of the two parts of a double-humped burst is difficult, because the second part is often too small to measure accurately, and the decay rate of the first part is modified by the appearance of the second part. Such measurements as could be made on decay-constant ratios were inconclusive.

#### *(c) The Relative Times of Arrival of Unpolarized Bursts on Different Frequencies*

In an earlier communication from this Laboratory(9) evidence was given of time-delays between the arrival of corresponding bursts on different frequencies. The bursts usually arrived in the order of decreasing frequency, the time interval for frequencies of 75 and 60 Mc/s. being of the order of a second. It was also stated that longer time differences had sometimes been observed, and a particular example was quoted of an outburst of very high intensity in which similar delays, but of the order of minutes, appeared to occur.

\* The bursts used in the timing measurements included a larger proportion of single bursts and a smaller proportion of complex bursts than occur in the whole burst population, owing to the relative ease in timing the former.



The present section takes up again the question of time-delays of the order of seconds, confirming and elaborating the original suggestion. However, on the question of longer delays, the present author has never since, in the recording of hundreds of bursts, obtained any evidence for delays of the order of minutes. Either the case reported earlier was very unusual, or the record was misinterpreted; as the relative amplitude of different portions of a complex burst may be very different on different frequencies, such a misinterpretation of a single case is quite possible.

In the measurements analysed here, only those bursts were used that were sufficiently similar on the frequencies concerned for there to be no reasonable doubt as to the corresponding portions. Also, as the rate of rise appears to vary with frequency, we used only the start of bursts rising from the normal quiet level; no use was made of the time of peaks or of succeeding portions of overlapping bursts. As the initial portion of bursts is approximately parabolic, the time of starting of each burst was marked by fitting to it one of a series of parabolas. Recorders were interchanged at intervals, and checked by observing simultaneous transients on different frequencies. It is considered that the differences between 85, 65, and 60 Mc/s. measured by this method are correct to  $\pm 0.25$  sec. Those between 60 and 19 Mc/s., in which  $\frac{3}{4}$  in./min. records were used, are correct to  $\pm 2$  sec.

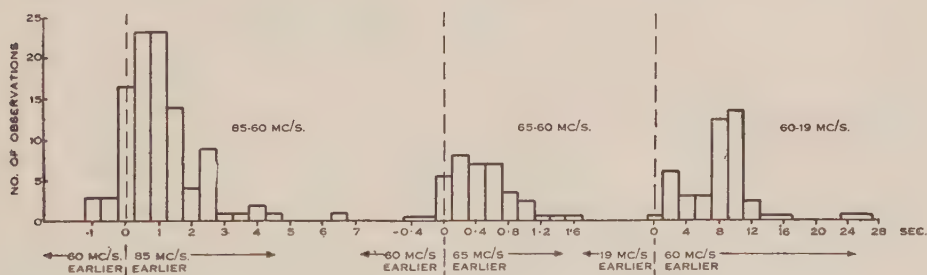


Fig. 7.—Relative times of arrival of corresponding bursts on different frequencies.

The results obtained by this method are set out in the histograms of Figure 7. The histogram for 85 and 60 Mc/s. has a fairly symmetrical distribution, with the peak definitely to one side of the zero, showing that the majority of 60 Mc/s. bursts occur between 0 and 1.5 sec. later than the corresponding 85 Mc/s. bursts, the most common time difference being about 0.7 sec. A similar histogram for delays between 65 and 60 Mc/s. shows a most common time difference of 0.3 sec. The histogram for 60 and 19 Mc/s. similarly shows the lower frequency arriving later, the most common delay being about 9 sec. The histograms show a tendency to skewness with a longer tail on the long delay side. Such a curve tends to occur when the quantity observed always has the same sign but shows considerable deviation about the mean value.

The scatter diagram of Figure 6 (B) shows that for those bursts in which 60 Mc/s. arrives before 85 Mc/s. the damping constant on 60 Mc/s. is always

greater than that on 85 Mc/s., though for the more common bursts on which 85 Mc/s. arrives first the two damping constants are on the average equal.

(d) *Association between Outbursts and Radio Fade-outs*

It has been observed several times that radio fade-outs are frequently accompanied by a large increase in solar noise and some statistical results have been given for 2700 Mc/s.(7) and 200 Mc/s.(4), the former showing a very good correlation and the latter a rather poor one. So far there has been no statistical information available on low frequencies.

Simultaneously with the recording on 85 and 60 Mc/s., we have monitored the carrier from VLQ3, a 10-kilowatt station operating on 14.4 Mc/s. about 450 miles north-east of our station. We also receive radio fade-out data from the ionospheric recorder at the Commonwealth Observatory, Mt. Stromlo.

TABLE 1

Source of Fade-out Data	Total No. of Fade-outs	Number of Related Occurrences of					
		Outbursts	Other Unpolarized Bursts			Enhanced Level	No Activity
			Large	Medium	Small		
Stromlo and VLQ3	18	11	2	0	0	2	3
Stromlo but not VLQ3 .. ..	15	0	2	1	2	1	9

Table 1 summarizes the results for January 1-September 23, 1948. A burst is considered to be associated with a fade-out only when it occurs close to it in time and is either the only appreciable activity recorded for some hours or is of outstanding magnitude compared with any other activity. In all except one case the associated solar noise occurred within 15 minutes of the fade-out. The exception is commented on later, and was of such high magnitude that there is little doubt it was associated with the fade-out.

The larger fade-outs, which are the only ones apparent on VLQ3, are almost inevitably accompanied by increased solar activity, more than half of them by an outburst. However, almost all the minor fade-outs that did not affect VLQ3 were unaccompanied by increased solar activity.

It is not possible to give an exact statement of the relative times of commencement of fade-out and outburst, as the fade-out usually builds up over a few minutes and the main outburst is frequently preceded by small bursts (see Fig. 2). Usually the outburst is not more than a few minutes after the fade-out. However, on January 24 a fade-out occurred at 1210 E.A.S.T., while the associated

solar noise did not commence till 40 minutes later and then continued at a very high level for about two hours. It is not certain whether the activity was an outburst or enhanced level, as the station was unattended.

In addition to the 11 outbursts which were correlated with fade-outs, there were nine further outbursts at times when no fade-out was recorded. Six of these occurred before 1100 local time, so that any accompanying flares would have been less likely to produce fade-outs than with the other 11 outbursts, only 2 of which occurred before 1100. Unfortunately, we have no data on flares over this period except those which can be inferred from fade-out data.

## V. DISCUSSION OF RESULTS

The observations recorded here have shown the existence of two distinct types of radiation from the sun other than the fairly steady thermal level. One of these we refer to as "enhanced radiation" and the other as "unpolarized bursts". Both are very variable and reach much higher intensities than the thermal level.

"Enhanced radiation", rare on the frequencies considered here, lasts for hours or days and is usually associated with the passage of a large spot group. It is normally circularly polarized, suggesting that the magnetic field of the spot group plays a part in its production. This is borne out by its close association with large sunspot groups. The spot field may not normally extend far enough into the corona to give rise to this type of radiation at 19 Mc/s.

The "unpolarized bursts", more common at these frequencies, are a transitory phenomenon, also associated with the passage of sunspot groups, but less markedly so. They would appear to arise outside the influence of sunspot magnetic fields, as they are not circularly polarized.

The "unpolarized bursts" decay exponentially, and Westfold(16) has suggested that the decay constant can be identified with the collision frequency. The observed values for 85 and 60 Mc/s. would then indicate that bursts at these frequencies arise at heights of about  $5 \times 10^5$  km. in the solar atmosphere. This is well out in the solar corona and past the region at which radiation at these frequencies directed into the sun would be turned back. (We shall refer to this region at which radiation would be turned back as the echoing region.) The lower values of decay constant at 19 Mc/s. suggest that the bursts observed on this frequency arise still further out in the corona. The range of decay constants at a single frequency may be due either to variations in position of the region of origin or to local variations in density and temperature.

The production of bursts outside the echoing region offers an explanation for the common occurrence of double-humped bursts, the second part being an echo of the original burst. The observed faster decay rate of single bursts will then be expected, as they will arise inside the echoing region where the collision frequency is higher than it is further out. Jaeger and Westfold(17) have calculated that for bursts on 60 or 85 Mc/s. originating out in the corona at about the echoing level for 20 Mc/s., the ratio of the intensity of echo to direct burst should be 1 to 4, agreeing with our observations, and that the delay between



echo and direct burst should be about 1 seconds, which is also in reasonable agreement with our observations.

All the observations discussed above suggest that bursts arise well out in the solar corona. Jaeger and Westfold have suggested that corresponding bursts on different frequencies have a *common* level of origin. According to their theory, the spectrum will have a peak in intensity at the frequency appropriate to its level of origin, falling off very steeply on the low-frequency side and more slowly on the high-frequency side. The relative intensity of bursts on 60 and 85 Mc/s. shows such a decrease with increasing frequency, but the evidence at 19 Mc/s. is inconclusive. On this theory all corresponding bursts might be expected to show the same decay rate. The slower rate observed on 19 Mc/s. is in disagreement with this.

The good correlation between major flares and outbursts and the poor correlation with minor flares suggest that both flares and outbursts are produced by the same kind of disturbance. When extensive enough, the disturbance may produce both a flare and large bursts, but when limited in action may produce only one or the other. It is not surprising that a very extensive disturbance should be needed to produce both a flare, which arises in the chromosphere, and an outburst, which at these frequencies we have shown to arise well out in the corona.

The observed delays between corresponding bursts on different frequencies may be due to their different group velocities in the solar atmosphere. The data given by Jaeger and Westfold(17) indicate that this mechanism could cause delays of up to 0.5 second between 85 and 60 Mc/s. and up to 1 second between 60 and 19 Mc/s., the lower frequency coming last in each case. Comparison with the observed values of about 0.7 second and 9 seconds suggests that this process is not the only cause of delays, particularly at the lower frequencies. Alternatively the delay may be due to the finite velocity of the disturbance initiating the burst, which may produce its maximum effect at different levels for different frequencies.

## VI. ACKNOWLEDGMENT

The work described in this paper was carried out as part of the research programme of the Division of Radiophysics, C.S.I.R.O.

## VII. REFERENCES

- (1) PAWSEY, J. L., PAYNE-SCOTT, RUBY, and MCCREADY, L. L.—*Nature* **157** : 158-9 (1946).
- (2) MCCREADY, L. L., PAWSEY, J. L., and PAYNE-SCOTT, RUBY.—*Proc. Roy. Soc. A* **190** : 357-75 (1947).
- (3) PAWSEY, J. L.—*Ibid.* **158** : 633-4 (1946).
- (4) ALLEN, C. W.—*Mon. Not. Roy. Astr. Soc.* **107** : 386-96 (1947).
- (5) APPLETON, E. V., and HEY, J. S.—*Phil. Mag.* **37** : 73-84 (1946).
- (6) LOVELL, A. C. B., and BANWELL, C. J.—*Nature*, **158** : 517-8 (1946).
- (7) COVINGTON, A. E.—*Proc. Inst. Radio Engrs. N.Y.* **36** : 454-7 (1948).
- (8) RYLE, M., and VONBERG, D. D.—*Nature* **158** : 339-40 (1946).



- (9) PAYNE-SCOTT, R., YABSLEY, D. E., and BOLTON, J. G.—*Ibid.* **160**: 256 (1947).
- (10) MARTYN, D. F.—*Ibid.* **158**: 308 (1946).
- (11) APPLETON, E. V., and HEY, J. S.—*Ibid.* **158**: 339 (1946).
- (12) RYLE, M., and VONBERG, D. D.—*Ibid.* **158**: 339-40 (1946).
- (13) PAWSEY, J. L.—*J. Instn. Elect. Engrs.* (in press).
- (14) RYLE, M., and VONBERG, D. D.—*Proc. Roy. Soc. A* **193**: 98-120 (1948).
- (15) WILLIAMS, S. E.—*Nature* **162**: 108 (1948).
- (16) WESTFOLD, K. C.—*Aust. J. Sci. Res. A* **2**: 169.
- (17) JAEGER, J. C., and WESTFOLD, K. C.—*Aust. J. Sci. Res. A* **2** (in press).

# THE NOISE-LIKE CHARACTER OF SOLAR RADIATION AT METRE WAVELENGTHS

By RUBY PAYNE-SCOTT\*

[*Manuscript received November 25, 1948*]

## *Summary*

Thermal radiation received from the sun at radio frequencies is indistinguishable from fluctuation noise in the radio receiver. However, the enhanced radiation often encountered at metre wavelengths is unlikely to have a thermal origin and hence may not be noise-like in character. This paper describes an experiment in which enhanced solar radiation on 85 and 60 Mc/s. was compared with receiver noise and with a C.W. signal by comparing the outputs from a diode with variable bias. This should detect any marked difference in the distribution of amplitudes of the fluctuations in solar and receiver noise. It was found that the enhanced solar radiation was indistinguishable from receiver noise.

## I. INTRODUCTION

The random motion of electrons in the components of a radio receiver produces a fluctuating output voltage. Such fluctuations are indistinguishable from those produced when radiation from a hot body is incident on an aerial connected to the receiver, and the name "thermal noise" has been applied to all fluctuations of this type. The term "solar noise" has been applied to the recently discovered radiation from the sun at radio frequencies, as the video output when monitored by a loud-speaker sounds like thermal noise. It is generally believed that the steady background level of solar radiation is thermal in origin, so that there is good reason for expecting this radiation to be identical with thermal noise.

However, during the passage of a large sunspot the radiation at metre wavelengths often rises to intensities thousands of times the normal level, and varies rapidly with time. When observed on a cathode-ray oscillograph or on headphones this enhanced radiation still resembles thermal noise(1), but it seems unlikely that it is thermal in origin. The aim of the experiment described here was to compare more precisely the fluctuations in enhanced solar radiation with those of thermal noise.

The statistical distribution of the amplitudes of the voltage fluctuations produced by thermal noise in a radio receiver has been studied by a number of workers, notably by Rice(2). The form of the distribution is independent of the receiver selectivity characteristics, and is approximately normal for the unrectified signal. A convenient criterion for comparing the distribution of amplitudes in voltages made up of varying proportions of thermal noise and

\* Division of Radiophysics, C.S.I.R.O.

C.W. signal is the rate of decrease in the rectified output as the bias on a detector is increased. This was adopted as a means for comparing the amplitude detection in solar noise with that of thermal noise with and without a superposed C.W. signal.

## II. EXPERIMENTAL

The experimental arrangement used was as follows :

Radiation at either 85 or 60 Mc/s. was fed into a conventional radar receiver with a bandwidth of about  $1\frac{1}{2}$  Mc/s. The receiver gain was so adjusted that the distortion of noise peaks was not serious. The diode detector was provided with a variable bias through a low-impedance source (Fig. 1). Simultaneous readings of diode current and bias voltage were taken for values of external

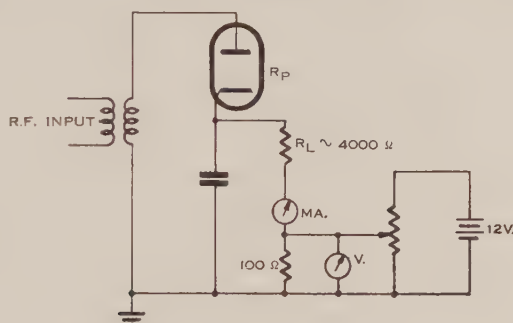


Fig. 1.—Detector circuit.

bias ranging from zero to a bias large enough to reduce the diode current practically to zero. Observations were made for each of the following arrangements :

(i) Receiver connected to aerial receiving solar noise. The receiver output is then due to a combination of receiver noise, cosmic noise, and solar noise. Let  $M$  be the total output with no external bias, and  $M_0$  the corresponding output in the absence of enhanced solar radiation, i.e. the output due to receiver noise and cosmic noise (determined with the aerial looking at the sun on previous "quiet" days).

(ii) Receiver connected to dummy load and gain increased so that output with zero external bias is again  $M$ . The input is now entirely receiver noise. This is analogous to condition (i) if solar noise is identical in character with receiver noise.

(iii) Receiver connected to signal generator of same impedance as aerial. The gain is decreased till the output with no external bias and no input signal is  $M_0$ . A C.W. signal at centre frequency is now injected and its intensity adjusted till the reading is again  $M$ . This is analogous to condition (i) if solar noise is a single-frequency signal.

Observations were made on August 5 and 6, 1948, when a very high level of solar radiation was received, associated with the passage of a large sunspot group. Both 85 and 60 Mc/s. were received by a horizontally-polarized Yagi aerial. The radiation was of the type referred to as enhanced radiation in the

preceding paper (see 3, p. 217). The level of the radiation was by no means steady, so that a number of readings in rapid succession were taken. The points obtained in a typical series of observations are shown in Figure 2. In this case the presence of solar radiation increased the detector current from 0.16 to 0.62 mA., so that the power received from the sun was about fourteen times that from the receiver and the sky together. The distribution of amplitudes with this mixture of solar radiation and receiver noise coincides with those for receiver noise alone within the accuracy of measurement. This was true for all the observations on both frequencies.

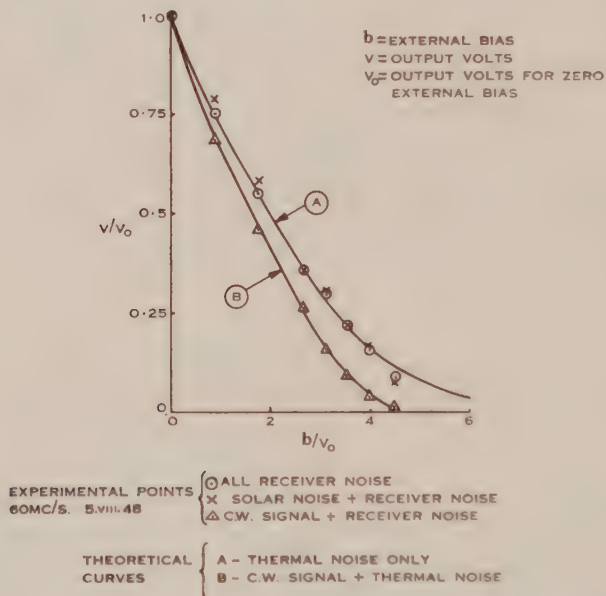


Fig. 2.—Rate of change of output with detector bias.

The curves in Figure 2 are theoretical, calculated from equation (4.3-15) given by Rice(2) for the output of a biased linear detector into which is fed a noise voltage together with a C.W. signal at centre frequency. In curve A the magnitude of the C.W. signal is zero, and in curve B it has the appropriate value for the mixture of noise and C.W. in the experimental case shown. In computing the curves, it is assumed that the detector is linear even for very small inputs and that its plate resistance is independent of voltage. Also the effect of varying detector bias on the A.C. impedance of the diode and hence on the gain on the preceding stage of the amplifier, has been neglected. Both these effects of varying diode impedance are small and tend to cancel one another. The observed points fit the theoretical curves well.

The results show that to within the limitations of this experimental technique there is no observable difference between solar noise and thermal noise over the small frequency range accepted by the receiver and that solar radiation certainly does not reach us as a series of discrete frequencies separated at intervals of more than, say, 2 Mc/s.



## III. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Radiophysics, C.S.I.R.O. The suggestion that the distribution of solar noise could be investigated by using a biased detector was made to us by R. E. Burgess of the National Physical Laboratory, England.

## IV. REFERENCES

- (1) PAWSEY, J. L., PAYNE-SCOTT, RUBY, and MCCREADY, L. L.—*Nature* **157** : 158-9 (1946).
- (2) RICE, S. O.—*Bell Syst. Tech. J.* **24** : 46-156 (1945).
- (3) PAYNE-SCOTT, RUBY.—*Aust. J. Sci. Res. A* **2** : 214 (1949).

# THE ARTIFICIAL STIMULATION OF PRECIPITATION BY MEANS OF DRY ICE

By P. SQUIRES\* and E. J. SMITH\*

[*Manuscript received November 30, 1948*]

## *Summary*

In experiments carried out near Sydney, supercooled water clouds were "seeded" with ice crystals by dropping pellets of solid carbon dioxide into them. Up to August 25, 1948, 20 experiments were made under conditions sufficiently well defined for the experiments to be significant. In 15 of them precipitation is believed to have been released artificially. Of 11 clouds with tops colder than  $-7^{\circ}\text{C}$ ., 10 precipitated. The depth of the treated clouds ranged from 1000 to 15,000 ft.

In none of these did natural precipitation occur within 20 miles. Further evidence that the effect is genuine is given by the fact that both the likelihood of success in inducing precipitation and the time between treatment of the cloud and the appearance of precipitation at its base, varied consistently with the cloud characteristics.

The observations are consistent with the view, as postulated by Bergeron, that the precipitation grew from ice crystals (formed by the dry ice pellets).

## I. INTRODUCTION

### *(a) Bergeron's Ice Crystal Theory of Precipitation*

The release of precipitation from clouds has been discussed by Bergeron(1). He considered several mechanisms by which drops the size of raindrops could originate in a cloud: direct condensation on to the initial cloud droplets, distillation from smaller or warmer droplets onto larger or cooler ones, and coagulation promoted by electrostatic or hydrodynamic forces. His conclusion was that, except in certain unusual circumstances, none of these processes is adequate to explain the initial release of "real precipitation"; some of them would at best contribute to the growth of raindrops, once formed, as they fell through the cloud. (The terms "real precipitation" and "real rain" as used by Bergeron were intended to exclude drizzle, such as commonly occurs in the warm sector of a temperate cyclone, falling from non-freezing stratiform cloud.) He then propounded his well-known theory that "almost every real raindrop ( $d > 0.5\text{ mm.}$ ) and all snow-flakes originated around an ice crystal". According to this theory, the initial release of precipitation occurs when ice crystals form in a supercooled water cloud. At temperatures above  $-10^{\circ}\text{C}$ ., water vapour usually condenses to water droplets; ice crystals often do not form until the temperature falls considerably lower. When ice crystals do form in a supercooled water cloud they are surrounded by an atmosphere supersaturated with respect to ice, so that water vapour sublimates onto them. As they grow the ambient vapour

\* Division of Radiophysics, C.S.I.R.O.

pressure falls below its initial value, that in equilibrium with supercooled water, and the drops evaporate. Making several simplifying assumptions, Bergeron deduced that, in a cloud supercooled to  $-10^{\circ}\text{C}$ . and containing 1000 droplets of diameter  $20\text{ }\mu$  per cc., an ice crystal infection of one per cc. would cause all the liquid water to evaporate and sublime onto the ice crystal in about 10–20 minutes.

It will be seen that artificial precipitation was released by this process in experiments described below.

### *(b) Chamber Experiments on Supercooled Water Fogs*

Chamber experiments on the “seeding” of fogs of supercooled water droplets with ice crystals were made by Schaefer(2). The ice crystals were formed by dropping pellets of “dry ice” or by inserting a rod cooled in liquid air. The ice crystals grew and fell out, and the supercooled fog evaporated, the vapour subliming onto the ice crystals. This took place in a period of the order of a minute, depending on the temperature and density of the fog and the intensity of the ice crystal infection.

Schaefer's chamber experiments have been confirmed by A. F. A. Harper of the Division of Physics, C.S.I.R.O., the results being personally communicated to the authors. Harper found that when dry ice pellets were dropped through a chamber containing supercooled fog, the fog was precipitated as ice crystals if its temperature was below  $-0.7 \pm 0.1^{\circ}\text{C}$ .

### *(c) Experiments on Natural Supercooled Water Clouds*

In the United States, Schaefer(2), and Langmuir *et al.*(3) treated supercooled clouds with dry ice with the object of releasing precipitation artificially. Numerous reports of these tests appeared in the Press. Coons, Gentry, and Gunn(4) have reported the results of 38 tests made on stratiform winter cloud in the United States. Precipitation was induced in seven of these tests, but in every case there was natural precipitation falling within 30 miles.

This paper presents the results of a programme of systematic flight tests to study the effect of treating natural clouds with dry ice, which was commenced early in 1947 by the Division of Radiophysics, C.S.I.R.O., in cooperation with the Royal Australian Air Force. Descriptions of some individual trials have already been reported by Kraus and Squires(5), and by Smith(6).

## II. DESCRIPTION OF EXPERIMENTS

### *(a) General*

In most cases the dry ice was dropped and the results observed from a single aircraft. Officers of C.S.I.R.O. directed the experiments and operated any special equipment installed. The aircraft operated from Richmond or Schofields, two R.A.A.F. aerodromes about 30 miles north-west of Sydney; on most of the flights they were operated by members of a special detachment of the Aircraft Research and Development Unit of the R.A.A.F. Additional observations were made by radar sets, and synoptic and radiosonde data were provided by courtesy of the Commonwealth Meteorological Bureau.

The uncertainty of meteorological forecasting rendered it impossible to adhere to a detailed prearranged plan. A flight was arranged for a given time, and details of the experiment were decided in the air to suit whatever clouds were encountered. The aircraft took off, climbed over base, and set off in the direction of the most promising clouds visible, whether over land or sea. When no clouds suitable for experiment were available locally, the synoptic situation was used to suggest the most favourable direction in which to fly in search of better conditions. It was occasionally necessary to fly as far as 300 miles from base.

### *(b) Choice of Clouds*

A large proportion of the tests was made on cumulus. Deep clouds were considered the most likely to yield heavy precipitation, and cumulus is the type of deep cloud most often encountered over New South Wales. With deep stratiform cloud it is difficult to relate the place where dry ice is dropped to that where any precipitation emerges from the bottom, but cumulus provides a deep cloud of limited horizontal extent which makes the results of experiments easier to interpret. Some experiments were also carried out with shallow subfreezing layers of stratocumulus or altocumulus. The object was to investigate the initial production of ice crystals without complications, such as strong vertical currents, associated with deep clouds.

In all cases the cloud tops appeared to consist initially of supercooled water. Clouds were avoided if they showed any sign of containing ice crystals, usually indicated by the fuzzy edges typical of ice clouds or by optical phenomena such as haloes, parhelia, or iridescence. Early flights indicated that certain types of cloud were unlikely to yield precipitation or useful information and subsequently these were avoided as far as possible. They included clouds in which the tops were blown off by a strong horizontal wind shear. Ice crystals formed in the top of these clouds could not fall into the lower part, and the tops seldom lasted very long. Cumuli which grew rapidly and then quickly dissipated were also found unsuitable.

In some of the later tests, where a choice of clouds was available, attempts were sometimes made to select the combination of height, temperature, and depth about which previous tests gave least information, or alternatively, the combination which illustrated best some point under current consideration.

### *(c) Granulated Dry Ice used for Seeding*

In the earlier tests some hundreds of pounds of dry ice were used on each cloud, but later this was reduced by a factor of about five, a typical quantity being 60 pounds per cloud. In recent flights it was dropped at a rate of from 10 to 30 pounds per mile from a chute fitted in place of one of the aircraft windows.

In the early tests, the dry ice was broken into fragments, the largest of which were such that the diameter of the sphere of equal volume (the "effective diameter") was about 2 cm. The whole of the resultant mixture, containing some 20 per cent. of dust, was used. Later, the material was sieved to limit the size of the largest pellet to an effective diameter of about 1.5 cm. and to exclude the dust, in order to secure a more reproducible method of infecting the cloud,



and to prevent the caking of the granulated material which occurs when the dust is included. These changes of practice did not produce any noticeable change in the resulting behaviour of treated clouds, but the variation of cloud type and behaviour is so wide that many more tests would be needed to make apparent any but the grossest of such effects. The whole series of tests is therefore treated for the present as homogeneous.

The constitution of the latest mixture is shown very approximately in Table 1. An estimate of the distance a pellet falls before evaporating, taken from values given by Langmuir(3) and referring to the mean diameter of the range, is given in the fourth line of the table. According to Langmuir these values have been found experimentally to be slightly too large. Approximate measurements by the authors in which a spherical pellet was supported by a vertical air current confirmed Langmuir's values within a factor of 2.

TABLE 1  
PELLET SIZES OF DRY ICE USED

Diameter of sphere of equal volume (cm.) . . . . .	0-0.2	0.2-0.4	0.4-0.6	0.6-0.8	0.8-1.0	1.0-1.2	1.2-1.4	1.4-1.6
Mass in each category (%) ..	1	4	11	21	22	20	13	8
Numbers in each category (%) ..	63	16	9	6	3	2	1	0.3
Distance fallen before evaporat- ing (ft.) ..	270	2300	5100	8500	12,000	15,000	21,000	26,000

In most of the experiments the dry ice was dropped in two runs over the cloud being seeded. The runs were made approximately at right angles, to achieve a more uniform distribution of ice crystals. The aircraft usually flew about 200 ft. above the top of the cloud during the drop. Thus all but the smallest particles of dry ice should have reached the top of the cloud, and some of the larger pellets should have reached the freezing level in even the deepest clouds.

#### (d) Scope and Plan of Observations

The scope of the observations varied considerably. Temperatures and sometimes humidity were measured in the clear air. Radiosonde observations taken at Rathmines, 70 miles NNE. of Sydney, were used to help in estimating conditions in the atmosphere. A 10-cm. aircraft radar, a 10-cm. ground radar at Richmond, and 10-cm. and 25-cm. radar sets at George's Heights, near the entrance to Sydney Harbour, were used on various occasions to observe rain areas. Visual and photographic observations were made from the aircraft

of the treated cloud and of precipitation. The intensity of the latter was estimated by flying through it and observing the rain or snow as it struck the aircraft, and also from its appearance and the extent to which it reduced the visibility. When it fell in populated areas, ground parties questioned local residents. All temperatures quoted below for clouds are those measured in the clear air at the same level, and all heights are those indicated on the aircraft altimeter. Optical phenomena were observed and photographed.

The plan of observations was as follows: Preliminary measurements were made of heights, temperatures, etc., the dry ice was dropped, and then the aircraft waited, usually at the bottom of the cloud. As far as the aircraft's endurance permitted, the observations were carried on until any effects had stopped, or, if nothing was seen, as long as seemed worth while in the light of past experience. In some of the early trials clouds may have been abandoned too soon.

### III. RESULTS OF EXPERIMENTS

#### *(a) The Convention Adopted in Defining "Precipitation"*

In these experiments it was found necessary to adopt a definition of the least intensity of precipitation of which note should be taken. Very light precipitation is difficult to detect unless the aircraft actually flies through it, which cannot be arranged unless the position of the precipitation is already known. In clouds of considerable horizontal extent it is difficult to relate the position where dry ice was dropped to the position where precipitation, if any, will emerge from the bottom. Also, if only a faint trace of precipitation is observed it is very difficult to make sure that natural precipitation of similar intensity is not falling from untreated clouds in the neighbourhood.

Consequently the following convention was adopted: note was taken of precipitation (induced or natural) only if it was visible as streaks at a distance of not less than 10 miles.

#### *(b) The Number of Significant Experiments*

Forty-one flights were made up to August 25, 1948. On some of them no clouds were found suitable for treating with dry ice; on a few, two or three clouds were treated. Altogether, 38 experiments were made, in 20 of which observations were adequate and conditions sufficiently clear cut for the experiments to be regarded as significant. The remaining 18 were not considered significant, for example, because it was subsequently found that adjacent clouds were precipitating, or because a confusing arrangement of cloud layers made it difficult to decide which part of the cloud lay below the place where the cloud top had been treated. In these inconclusive experiments there appeared to be nothing which was not consistent with the conclusions stated below.

#### *(c) Experiments in which Precipitation is Believed to have been Released Artificially*

In the nature of these experiments, it is never certain whether or not a cloud which is treated with dry ice and begins soon afterwards to precipitate

would have done so without artificial aid. However, in the 15 cases where it is believed that precipitation was induced, other similar cloud in the vicinity remained unaffected. Where there was the slightest doubt, because other precipitation occurred in the neighbourhood, the experiment was rejected as inconclusive. While the aircraft was flying around the cloud into which dry ice had been dropped, a careful watch was kept on adjacent clouds. The ground radar observations were used when available to check that there was no natural rain in the neighbourhood of the treated cloud. In no case listed below as a positive result did natural precipitation as defined above occur within 20 miles of the treated cloud. In all but one case no natural precipitation was observed within 30 miles. A summary of the experiments in which it is believed that precipitation was released is given in Table 2.

It will be seen that in most of the positive tests the cloud top extended up to a temperature below  $-7^{\circ}\text{C}$ . The clouds were stable, persisting for considerable periods, and were solid and compact; for instance, their tops were not being blown off by a horizontal wind shear. Of these flights seven were made in winter, four in summer, and four in autumn, and there is nothing in the results to show that the season had any influence on the chance of success. In three cases the precipitation occurred over the sea; in 12 over the land.

None of the positive results occurred in purely continental air which, over coastal New South Wales at least, is very often cloud-free. In several of the positive summer tests, however, there was only a shallow layer of tropical maritime air (usually less than 5000 ft. deep) below a continental air mass. Several of the positive winter tests occurred in modified maritime air streams from the south-west, with a previous land trajectory of some 500–800 miles. Only two of the positive results occurred in purely maritime air masses.

(i) *Cases where induced Precipitation reached the Ground.*—In 10 of the 15 cases where precipitation was induced it was seen, from the aircraft, to reach the ground. Owing to the sparseness of population it was not always possible to confirm this by interviewing local residents. In these 10 cases, the thickness of the clouds was greater than 3000 ft. and the bases were within 8000 ft. of the ground or sea, but as the tops of the clouds which were induced to yield precipitation all lay in the restricted range between freezing level and 7000 ft. above it, the thickness and height of base were not unrelated.

(ii) *Cases where Precipitation was Induced but did not reach the Ground.*—In three of the 15 cases where precipitation was induced it was seen to evaporate before reaching the ground. In two other cases this point was not definitely determined. In the former three, the thickness was not more than 2000 ft. and the height of the cloud base was more than 9000 ft. above the ground.

(iii) *Time elapsed to Commencement of Precipitation at Base.*—It was not possible to observe precisely the moment at which the first precipitation fell out of the cloud base. Presumably the first elements were too small and too few to be visible. Hence, in conformity with the convention regarding the definition of "precipitation" mentioned above, the "first precipitation" is defined as the first which was visible as streaks at a range of 10 miles. In nearly all cases



TABLE 2  
SUMMARY OF POSITIVE TESTS

Total Cloud Depth (ft.)	Height of Base Above Ground (ft.)	Precipitation Com- menced (min. after first seeding)		Type and Max. Intensity	Max. Area (sq. miles)	Precipita- tion Stopped (min. after first seeding)	End of Observa- tion (min. after first seeding)	Did Rain Reach Ground ?
		Observed Visually	Observed by Radar					
12000	8000	—	18	—	—	—	—	?
12000	8000	<25	16	Medium rain	> 20	> 100	c. 100	Yes
15000	3500	—		—	—	—	—	Yes
2500	5000	<10		Light snow	c. 0.1	—	12	?
11000	6000	22		Light rain	1	57	62	Yes
11000	7000	<20	18	Light rain	4	c. 46	50	Yes
7000	5000	19		Medium rain	30	> 59	59	Yes
5300	5500	22<<43		Light rain	6	> 68	68	Yes
2000	9400	16		Medium snow	2	28	60	No
1500	11000	12		Light snow	6	> 29	29	No
1000	9600	9		Very light snow	2	c. 30	56	No
5000	3000	60	34	Very light rain	> 0.5	> 100	100	Yes
3700	6700	14		Light snow	2	<41	55	Yes
3900	1800	12		Medium snow	1	32	38	Yes
3900	1800	12		Medium snow	16	57	72	Yes

as; Sc., Stratocumulus.



the precipitation was first seen from a range of several miles, and the observed interval was taken uncorrected; but in a few, when the clouds were of small extent, it was first seen at very close range. The greatest correction applied—which was added to the actual observed time—was three minutes. As the apparent density of the precipitation streaks always increased rapidly this did not introduce an uncertainty of more than a minute or two.

The appearance of precipitation at the base was observed visually in nine cases. In four cases, rain echoes were observed by radar. The intervals found in this way were not directly comparable with those found by visual observation of the cloud base, because the radar would sometimes detect rain before it reached the base of the cloud and because the threshold of "radar visibility" did not necessarily coincide with the convention "visible at a range of 10 miles" described above. In one case the commencement of precipitation was not observed either visually or by radar.

(iv) *Ice-cloud Optical Phenomena*.—In three experiments described elsewhere in detail by Smith(6), optical phenomena typical of ice clouds—haloes, parhelia, and parhelic circles—were observed after the treatment. On all these occasions the cloud had been examined before it was treated to make sure that no such phenomena could be seen.

(v) *Growth of a Cumulus into a Cumulonimbus*.—In the second experiment of February 5, 1947, described by Kraus and Squires(5), a towering cumulus grew into a cumulonimbus. The top of this cloud was initially at 23,000 ft., which was higher than in any of the other experiments and far higher than in most.

(d) *Experiments in which Precipitation Apparently was not Induced*

Data referring to experiments in which precipitation was *not* induced are given in Table 3. This includes only cases where the cloud remained reasonably "solid" during the period of observation; cases where the cloud evaporated quickly are excluded.

TABLE 3  
SUMMARY OF NEGATIVE TESTS

Date	Cloud Type	Temperature of Clear Air at Cloud Top (°C.)	Total Depth of Cloud (ft.)	Period During which Base of Cloud was Observed (min. after first seeding)
1948				
Feb. 19 .. ..	Cumulus	—6	9,500	23 to 59
Mar. 16 .. ..	Cumulus	—3	9,000	12 to 45
Apr. 1 .. ..	Cumulus	—22	14,000	25 to 77
Apr. 14 .. ..	Cumulus	—7	3,000	5 to 18, 25 to 50
June 7 .. ..	Cumulus	—2	6,000	7 to 30, 44 to 47

All these clouds were observed for at least 45 minutes ; the experience of the positive tests indicates that any effects usually appear within 25 minutes.

All but one of the failures occurred in clouds with tops warmer than  $-8^{\circ}\text{C}$ . The exception occurred on April 1, 1948, when the cloud treated was one of a line of large cumuli forming over the low ranges of the upper Hunter River valley, some 90 miles north of Sydney. The cumuli grew as they drifted south-east down the valley. This cloud was observed up to 77 minutes after seeding, by which time its neighbours in the line of cumuli were both raining heavily over a wide area and one of them had developed into a thunderstorm. It seems possible, therefore, that there was something abnormal in the cloud treated, since it is unlikely that the introduction of 85 pounds of dry ice would prevent a cumulus as large as this from raining. Its base was about 10 miles across, and its depth over two and a half miles, so that its volume must have been of the order of 50 cubic miles.

TABLE 4  
EFFECT OF CLOUD-TOP TEMPERATURE ON RESULT OF EXPERIMENT

Cloud-Top Temperature ( $^{\circ}\text{C}$ .)	Positive Results	Negative Results	Total	Positive Results (%)
0 to $-3$ ..	2	2	4	50
$-4$ to $-7$ ..	3	2	5	60
$-8$ to $-11$ ..	5	0	5	100
$-12$ to $-15$ ..	5	0	5	100
$-16$ to $-19$ ..	0	0	0	—
$-20$ to $-23$ ..	0	1	1	0
Totals ..	15	5	20	75

#### IV. DISCUSSION OF RESULTS

##### (a) *The Effect of Cloud Top Temperature on Results of the Experiment*

The chance of success in releasing precipitation varied with the temperature in the clear air level with the cloud top. The association between them is shown in Table 4.

From this table it can be seen that clouds with tops colder than  $-7^{\circ}\text{C}$ . almost always yielded precipitation when seeded with dry ice, but the chance of success was less when the cloud tops were warmer.

The apparently anomalous failure in the  $-20$  to  $-23^{\circ}\text{C}$ . range occurred on April 1, 1948, and, as described above, it seems possible that there was something unusual in the treated cloud. The association is shown in histogram form in Figure 1.

*(b) Effect of Depth of Cloud*

It will be seen from Tables 2 and 3 that the total cloud depth had no apparent influence on the chance of success, failure occurring with clouds of varying depth. The thinnest cloud treated, which was 1000 ft. deep, yielded a positive result. Presumably, however, if the treated cloud is too thin, any precipitation induced will consist of particles too small to form visible streaks.

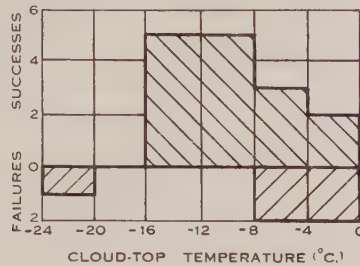


Fig. 1.—Distribution of successes and failures with cloud-top temperature.

*(c) Chance of Precipitation Reaching the Surface*

In these experiments the ratio of the total depth of cloud to the height of the cloud top above the ground is correlated with the chance of the precipitation reaching the ground. This is shown in Table 5 for all cases where it was definitely observed whether or not rain reached the ground.

Thus in this series of experiments, if a cloud had a depth more than about one-third of the height of its top above ground, any precipitation released reached the ground. This, of course, has no bearing on whether such a cloud can be

TABLE 5

AN ASSOCIATION WITH RAIN REACHING THE GROUND

Date	Depth of Cloud Height of Top	Did Rain Reach Ground ?
June 17, 1948 .. ..	0.09	No
June 11, 1948 .. ..	0.12	No
April 27, 1948 .. ..	0.18	No
August 17, 1948 .. ..	0.35	Only just
April 15, 1948 .. ..	0.49	Yes
April 13, 1948 .. ..	0.58	Yes
February 5, 1947 .. ..	0.60	Yes
March 10, 1948 .. ..	0.61	Yes
August 3, 1948 .. ..	0.63	Yes
January 28, 1948 .. ..	0.65	Yes
August 25, 1948 .. ..	0.69	Yes
August 25, 1948 .. ..	0.69	Yes
April 21, 1947 .. ..	0.81	Yes

artificially induced to precipitate in the first place. This ratio apparently formed a parameter roughly summing up the relevant factors, namely, the number and size of the precipitation elements which left the cloud base, and the depth, relative humidity, and temperature of the air below.

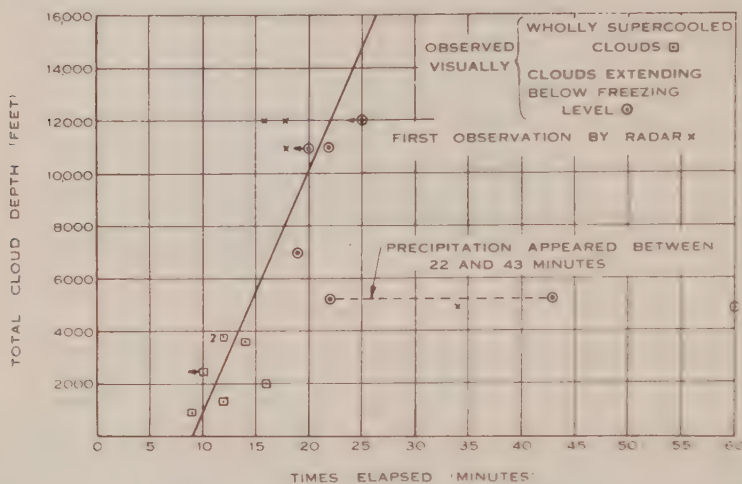


Fig. 2.—Time elapsed from first seeding to commencement of precipitation. (Arrows pointing to left indicate cases where the time elapsed was less than the value plotted.)

(d) *Time from Treatment of Cloud to Appearance of Precipitation at Base*

As described in Section III (b) above, the time of appearance of precipitation at the base was taken as that at which it became visible at a range of 10 miles. In Figure 2, these times are plotted against cloud depth. For very thin clouds the time taken is of the order of 10 minutes, tending to increase with increasing cloud depth. The scatter of the points may be due to the process being influenced by other factors such as cloud temperature; there are, however, insufficient data to permit these effects to be established. For instance, the three plots which lie well to the right of the general run of the points refer to two cases (April 15 and August 3, 1948, see Table 2) where the cloud-top temperature was higher than in the rest of the successful experiments ( $-2.5$  and  $-3.5^{\circ}\text{C}.$ ) respectively). The increase in time with depth is no doubt caused by the time taken for precipitation to fall to the bottom of deep clouds, but the slope of the graph cannot be taken as giving the average rate of fall, since the height at which the precipitation originated is unknown.

(e) *The Genuineness of the Effect*

The following evidence shows that there is a genuine causal relationship between the treatment of the supercooled clouds with dry ice and the release of precipitation.



(i) As described in Section III (b) above, 15 clear positive results were obtained on occasions when no natural precipitation whatever was observed within a radius of 20 miles from the treated cloud.

(ii) There is a distinct correlation between the temperature of the cloud top and the chance of success.

(iii) The times elapsed from treatment to the first appearance of precipitation at the cloud base nearly all fall within the restricted range between 10 and 20 minutes and, in addition, when plotted against total depth show a considerable degree of coherence. This could hardly occur if the release of precipitation were unrelated to the treatment.

#### V. THE RELATION OF THE RESULTS TO BERGERON'S ICE CRYSTAL THEORY

The observed precipitation must have formed, in the first instance, either as ice or as water. Arguments presented below lead to the conclusion that the precipitation elements grew from ice crystals formed by the dry ice pellets in the subfreezing part of the cloud. The mechanism involved in this case differs from that postulated by Bergeron in only one respect—that the ice crystals were introduced artificially instead of forming naturally on sublimation nuclei.

##### *(a) The Effect of Dry Ice*

The only known relevant effect of dry ice falling through a cloud is that found in the chamber experiments described in Section I (a) above: the formation of a large number of ice crystals which, provided the cloud is colder than  $-0.7^{\circ}\text{C}.$ , survive, grow steadily, and fall. No observations have been made which indicate that the dry ice pellets could promote colloidal instability in any other way.

##### *(b) Snow and Ice Precipitation*

In seven of the 15 positive tests the clouds were wholly colder than freezing, and the precipitation was in the form of snow or ice crystals in all these cases. As mentioned in Section III (c) above, on three of these occasions optical phenomena typical of ice clouds were seen after clouds were treated, the treated portion of the clouds dissolving into a thin fog of ice crystals. In view of the chamber experiments described in Section I (b) it is scarcely conceivable that the precipitation was produced otherwise than from the ice crystals formed in the wake of the pellets. In these cases, then, all the precipitation grew from ice crystals. The points in Figure 2 which refer to these seven clouds (squares) are in good agreement with those referring to clouds extending below the freezing level (circles). The implication is that the same general processes were at work in all cases. Further, the addition of lower cloud layers extending below the freezing level will not influence the processes in the upper supercooled part. It follows that in these cases some at least of the precipitation almost certainly formed from ice crystals.

(c) *The Effect of Cloud Characteristics on Speed of Action*

The graph of Figure 2 shows that the time taken for precipitation to reach the base of the cloud tended to increase with the cloud's depth. This would be expected if the precipitation originated in the upper (i.e. colder) part of the cloud. Also, the fact that the two clouds with the warmest tops took particularly long to yield precipitation is consistent with the ice crystal theory, since the rate of growth of ice crystals is less at temperatures closer to freezing.

(d) *The Effect of Temperature on the Chance of Success*

The chance of success in releasing precipitation from a cloud tends to decrease as the temperature of the cloud top rises above about  $-8^{\circ}\text{C}$ . This is consistent with the ice crystal theory, since at temperatures nearer to freezing the ice crystals would take longer to grow and would have more chance of escaping from the subfreezing region of the cloud, which is, of course, shallower. Also there may be in natural clouds a critical temperature below  $0^{\circ}\text{C}$ . (the value of which is not known, and may vary) similar to that mentioned in Section I (b), above which ice crystals do not grow.

## VI. CONCLUSIONS

1. Precipitation can be released from some supercooled water clouds by treating them with dry ice.
2. Factors which favour success are :
  - (i) The cloud should be compact and remain so for a period of the order of half an hour, particularly in the subfreezing layers—it should not evaporate or shear off.
  - (ii) The temperature of its top should preferably be below about  $-7^{\circ}\text{C}$ . Given the requirements in (i), at such temperatures the chance of success is quite high—in this series of experiments, about 90 per cent.—but it decreases at higher temperatures.
  - (iii) Depth does not appear to be relevant, although the precipitation released from a very shallow cloud is likely to be light.
3. The chance of precipitation reaching the ground depends on a number of factors, but in this series of experiments—under the climatic conditions of the central coast of New South Wales—it did so whenever the cloud depth was greater than one-third of the height of the cloud top above the surface.
4. The time taken for precipitation to appear at the cloud base increased with the depth of the cloud.
5. When the treated cloud was wholly supercooled, all of the precipitation was formed in the manner postulated in Bergeron's ice crystal theory. When the treated cloud extended below the freezing level some at least of the precipitation was almost certainly formed in this way.

## VII. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Radiophysics, C.S.I.R.O. Grateful acknowledg-

ment is made to the Royal Australian Air Force, which provided and flew the aircraft used in these experiments.

#### VIII. REFERENCES

- (1) BERGERON, T.—*P.V. Mét. Un. Géod. Géophys. Int.* Lisbon (1933).
- (2) SCHAEFER, V. J.—*Science* **104**: 457-9 (1946).
- (3) LANGMUIR, I., *et al.*—Gen. Elec. Res. Lab. Quart. Prog. Rep. No. 1, July 1947.
- (4) COONS, R. D., GENTRY, R. C., and GUNN, R.—First partial report on the artificial production of precipitation: stratiform clouds, Ohio, 1948. *Bull. Amer. Met. Soc. Ohio* **29**: 266-9 (1948).
- (5) KRAUS, E. B., and SQUIRES, P.—*Nature* **159**: 489-92 (1947).
- (6) SMITH, E. J.—*Aust. J. Sci. Res. A* **2**: 78-91 (1949).

# THE IODINATION OF AROMATIC COMPOUNDS

## II. THE IODINATION OF AROMATIC ETHERS AND OF 2-NAPHTHOL

By L. JURD\*

[Manuscript received February 17, 1949]

### Summary

A rapid process for the iodination of aromatic ethers is described, the ether in 95 per cent. ethanol being treated with iodine and mercuric oxide at an elevated temperature. Anisole and the 2-alkoxy-naphthalenes give high yields of iodo- compounds by this process.

1-Iodo-2-naphthol is readily prepared by the action of iodine and alkali or iodine and mercuric oxide on 2-naphthol.

A mechanism is suggested for the iodination of phenols and ethers in the presence of mercuric oxide.

### I. INTRODUCTION

Although low temperatures and anhydrous solvents are usually employed in the direct iodination of aromatic ethers by means of iodine and mercuric oxide it is now found that in the presence of water and by the use of an elevated temperature the reaction period is reduced very considerably without decrease in the yield and purity of the product. The marked decrease in the iodination of the phenol ethers, the rates decreasing in the order  $\text{CH}_3\text{O}- > \text{C}_2\text{H}_5\text{O}- > \text{HOOC}(\text{CH}_2)_2\text{O}- > \text{C}_6\text{H}_5\text{CH}_2\text{O}-$ , is not nearly so apparent with the 2-naphthol ethers, all of which are readily iodinated (Table 1).

### II. EXPERIMENTAL

(i) *Preparation of p-Iodo-anisole.* Iodine (19.9 g.) is added in four portions during 10 minutes, with vigorous agitation between each addition, to a suspension of mercuric oxide (13.1 g.) in a 95% ethanol solution (60 cc.) of anisole (8.5 g.), the temperature of the reaction mixture being maintained at 50–60 °C. After shaking for a further five minutes the temperature is raised to the boiling point of alcohol and the insoluble mercury compounds are removed by filtration and washed with boiling alcohol (30 cc.). On cooling, the filtrate, which deposits a mass of crystals, is poured into water (350 cc.) containing potassium iodide (3.0 g.). The light yellow, crystalline precipitate (17.2 g.) which separates is collected and recrystallized from 85% ethanol, colourless plates, melting at 51 °C. (lit. 50–51 °C.), being obtained (15.2 g., i.e. 83% of the theoretical yield).

The iodo ethers which have been prepared by this method are set out in Table 1.

1-Iodo-2-naphthyl benzyl ether, colourless plates, melting at 86.7 °C. from alcohol, and 1-iodo-2-naphthyl ethyl ether, colourless elongated prisms, melting at 75 °C. from petroleum ether, have not been previously described. The position of attachment of the iodine was proved by the synthesis of each from 1-iodo-2-naphthol by alkylation in the usual manner.

\* Massey Agricultural College, Palmerston North, New Zealand.



(ii) *Iodination of 2-Naphthol*.—1-Iodo-2-naphthol is an unstable substance for which only one satisfactory synthesis has been described(1). Although Messinger and Vortmann(2) isolated a highly coloured, amorphous product of indefinite composition by treating 2-naphthol in aqueous alkaline solution at 50 °C. with iodine, crystalline 1-iodo-2-naphthol is readily obtained in almost quantitative yields when slightly less than the theoretical quantity of iodine reacts with 2-naphthol in ice-cold, alcoholic potassium hydroxide solution.

Iodine (3.48 g.) is added gradually during 15 minutes to an agitated, alcoholic solution (10 cc.) of 2-naphthol (2.0 g.) and potassium hydroxide (1.60 g.) maintained at 5 °C. The flocculent, off-white precipitate which separates when the yellow, alcoholic solution is poured into dilute, aqueous sulphuric acid (80 cc.) is collected, washed thoroughly with water, and recrystallized from acetic acid. Colourless needles, melting at 94 °C., are obtained (3.44 g., i.e. 94% of the theoretical yield).

TABLE 1

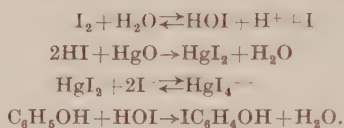
THE IODINATION OF AROMATIC ETHERS IN ALCOHOLIC SOLUTION AT 50–60 °C. IN THE PRESENCE OF MERCURIC OXIDE

Ether	Reaction Period (min.)	Product	Yield (%)	Melting Point (°C.)	Melting Point (lit. °C.)
Anisole .. ..	15	<i>p</i> -Iodo-anisole	83	51	50–1
Phenetole .. ..	40	<i>p</i> -Iodo-phenetole	67	28–9	29
Phenyl benzyl ether ..	60	<i>p</i> -Iodophenyl benzyl ether	Small	62	62
Veretrole .. ..	25	4-Iodo-veratrole	60	172–4/ 32 mm.	
Phenoxy-acetic acid	30	4-Iodo-phenoxy-acetic acid	62	156	155–6
2-Methoxy-naphthalene	20	1-Iodo-2-methoxy-naphthalene	89	88	88
2-Ethoxy-naphthalene	20	1-Iodo-2-ethoxy-naphthalene	82	75	
2 - Benzyloxy - naphthalene .. ..	20	1 - Iodo - 2 - benzyloxy - naphthalene	80	86–7	
2-Naphthoxy-acetic acid	30	1-Iodo-2-naphthoxy-acetic acid	71	174	d., 145, melts, 203–10

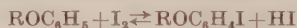
1-Iodo-2-naphthyl acetate, not previously described, prepared by the action of fused sodium acetate and acetic anhydride on 1-iodo-2-naphthol, is obtained from alcohol as small, colourless, brittle crystals, melting at 69 °C.

1-Iodo-2-naphthol is readily obtained by the following process: Iodine (3.50 g.) is added cautiously during 30 minutes to a mixture of mercuric oxide (2.35 g.) and 2-naphthol (2.0 g.) in alcohol at 5 °C., the reddish colour of iodine being allowed to disappear before each new addition is made. Vigorous agitation is continued for 10 minutes after completing the addition of iodine, the insoluble mercury compounds then being filtered off and washed with cold alcohol (5 cc.). The light yellow precipitate which separates when the alcoholic filtrate is poured into water (175 cc.) is collected, washed with dilute potassium iodide solution and water, and dried (3.70 g.). This product is dissolved in excess of glacial acetic acid with gentle warming. On cooling, tufts of colourless needles, melting at 93.5 °C. (lit. 93–4 °C.), separate (1.54 g.). On adding water to the warmed filtrate and cooling, long colourless needles, melting at 93 °C., separate (1.91 g.). The yield of 1-iodo-2-naphthol is therefore 92%.

In the presence of mercuric iodide the rate of reaction between iodine and 2-naphthol is accelerated considerably. The iodination of phenols in the presence of mercuric oxide therefore probably follows the course:



Since *p*-iodo-anisole and 1-iodo-2-methoxy-naphthalene are recovered quantitatively when equimolecular quantities of the iodo-ether and hydrogen iodide are agitated in alcoholic solution at 60 °C. for 30 minutes, the suggestion(3, 4) that the function of mercuric oxide in the iodination of aromatic ethers is to prevent the reversal of the reaction



is untenable. It is probable that the active agent is hypoiodous acid and that the role of mercuric oxide in the iodination of ethers is similar to that in the iodination of phenols.

### III. REFERENCES

- (1) VARMA, P. S., MOZUMDAR, D. W., and RAJAH, K. J.—*J. Indian Chem. Soc.* **10**: 595 (1933).
- (2) MESSINGER, J., and VORTMANN, G.—*Ber. dtsh. chem. Ges.* **22**: 2312 (1899).
- (3) BLICKE, F. F., and SMITH, F. D.—*J. Amer. Chem. Soc.* **50**: 1229 (1928).
- (4) HOCH, R., and CULBERTSON, J. B.—*Proc. Iowa Acad. Sci.* **47**: 265 (1940).

# ALKALOIDS OF THE AUSTRALIAN RUTACEAE: *MELICOPE FAREANA*

## I. ISOLATION OF THE CONSTITUENT ALKALOIDS

By J. R. PRICE\*

[Manuscript received March 9, 1949]

### Summary

The bark of the rain-forest tree *Melicope fareana* F. Muell. belonging to the family Rutaceae contains the four alkaloids *melicopine* ( $C_{17}H_{15}O_5N$ ), *melicopidine* ( $C_{17}H_{15}O_5N$ ), *melicopicine* ( $C_{18}H_{19}O_5N$ ), and *acronycidine* ( $C_{15}H_{15}O_5N$ ). The first three are each present to the extent of about 1 per cent. while the acronycidine content is about 0.1 per cent. About 1 per cent. of lupeol is also present in the bark. The leaves contain melicopine, melicopidine, melicopicine, and the known furanoquinoline alkaloid skimmianine. Melicopine and melicopidine contain two methoxyl groups, a methylimino and a methylenedioxy group. Melicopicine contains four methoxyl groups and a methylimino group.

## I. INTRODUCTION

This paper is the first of a series dealing with the detailed chemical examination of alkaloids occurring in native Australian plants. During World War II a number of such plants, chosen on the basis of botanical relationship to known drug producing species, or because of popular reputation(1), were investigated. The investigation is now being extended,† in the first instance to rain-forest plants belonging to the family Rutaceae(3). This family is an interesting one from the point of view of alkaloid content, since its members contain representatives of several of the principal structural classes, viz. quinoline, isoquinoline, carboline, and imidazole derivatives(4).

The present paper deals with the isolation of alkaloids from the bark and leaves of *Melicope fareana* F. Muell. The genus *Melicope*, according to Engler and Prantl(5), belongs to the family Rutaceae, tribe Xanthoxyleae, which includes such well-known alkaloid-bearing genera as *Xanthoxylum*, *Fagara*, and *Erodia*. The genus consists of at least 50 described species, eight of which occur in eastern Australia. *M. fareana* is restricted to the rain-forests of tropical

\* Division of Industrial Chemistry, C.S.I.R.O.

† The present programme involves collaboration of the C.S.I.R.O. with the Chemistry and Physiology Departments of several Australian universities. The search for alkaloids is being carried out in three stages. The first stage consists of systematic small-scale qualitative tests made in the field(2). Plants selected on the results of the field tests are then examined in the laboratory. Should the laboratory examination confirm the presence of tractable alkaloidal constituents, these alkaloids are isolated and their chemistry and pharmacology fully investigated.

Queensland. Although generally a small tree, it may reach a height of about 70 feet and a stem diameter of 12 inches. The inner bark, when cut, is a vivid yellow.\* The bark and leaves are rich in alkaloids which can be isolated in a crystalline condition without difficulty. The bark contains four alkaloids, none of which have been reported from other sources. These have been named *acronycidine*,  $C_{15}H_{15}O_5N$ , (isolated simultaneously from *Acronychia baueri* by Lahey and Thomas 6), *melicopine*,  $C_{17}H_{15}O_5N$ , *melicopidine*,  $C_{17}H_{15}O_5N$ , and *melicopicine*,  $C_{18}H_{19}O_5N$ . The total alkaloid content is about 3 per cent. of the dry weight of the bark, consisting of approximately 1 per cent. each of melicopine, melicopidine, and melicopicine, and 0.1 per cent. acronycidine, though the amount and proportions vary somewhat from tree to tree. The triterpene alcohol, lupeol, has been isolated from the bark in which it is present to the extent of about 1 per cent. It has previously been isolated from the bark of the related *Melicope erythrococca*(7) as well as from a number of other members of the Rutaceae, for example, *Xanthoxylum macrophyllum*(8). The leaves contain 1.5 per cent. of alkaloids from which were separated approximately 0.2 per cent. melicopine, 0.8 per cent. melicopidine, 0.03 per cent. melicopicine, and 0.3 per cent. of colourless base. This consisted almost entirely of the furanoquinoline alkaloid skimmianine which has been found previously in several rutaceous genera(9, 10). The skimmianine,  $C_{14}H_{13}O_4N$ , has been identified by the melting points and analyses of the base and its picrate, by its conversion to *isoskimmianine* on heating with methyl iodide, and by comparison with an authentic specimen.

The three major alkaloids, melicopine, melicopidine, and melicopicine, are yellow and nicely crystalline, giving strongly fluorescent solutions in organic solvents. They are optically inactive. All three are very weak bases, extracted by chloroform from aqueous acid solutions. Melicopidine, the strongest base of the three, has been characterized by its *picrate* and *picrolonate*, but melicopine and melicopicine do not form salts with these acids. Salts with mineral acids are unstable and have not been analysed. Because of this instability, it was found desirable to make as little use of acid as possible during the isolation. Melicopine and melicopidine contain two methoxyl groups and a methylimino group and give a positive methylenedioxy test. Melicopicine contains four methoxyl groups and a methylimino group, but the test for methylenedioxy is negative. The pharmacology of these three alkaloids was not investigated because they are too weakly basic to give water-soluble salts.

## II. EXPERIMENTAL

All melting points are corrected unless otherwise stated. Microanalyses were carried out by R. B. Bradbury and D. J. Clark.

### (a) Extraction of the Bark

The bark, milled to pass a 40 BS sieve, was extracted with methanol, either in a soxhlet extractor or by preliminary cold percolation followed by hot extraction. The latter procedure was employed when dealing with large quantities and the two extracts were worked up separately. The methanol solution was concentrated, the concentrate diluted with two-three times its volume

\* The popular name of the tree is "mustard bark".



of water, and extracted three times with chloroform. The dark brown viscous water layer after acidification and filtration from the tarry precipitate gave a precipitate with Mayer's reagent but only small amounts of resinous basic material could be isolated from it. The chloroform extracts were shaken with caustic soda (5%) to remove phenolic and acidic substances which contained only a negligible quantity of alkaloidal materials. The chloroform was then concentrated to a small volume and diluted with 8–10 volumes of ether, when a dull yellow crystalline solid separated. This precipitate was dissolved in a little chloroform and reprecipitated by ether. The combined ethereal filtrates were concentrated, the residue again dissolved in a little chloroform and retreated with ether. The precipitate was again dissolved, reprecipitated, and the solid added to the main precipitate.

(b) *Ether-Insoluble Fraction—Melicopine*

Recrystallization from ethanol of the solid precipitated by ether gave *melicopine* as yellow needles, m.p. 178.5–179.5 °C.

Found: C, 65.3; H, 5.0; N, 4.4;  $\text{CH}_3\text{N}$ , 7.9;  $\text{CH}_3\text{O}$ , 20.1%; mol. wt. (11), 312;  $[\alpha]_{\text{D}}$ , 0° in chloroform.

Calculated for  $\text{C}_{17}\text{H}_{15}\text{O}_5\text{N}$ : C, 65.2; H, 4.8; N, 4.5;  $\text{CH}_3\text{N}$ , 9.3;  $\text{CH}_3\text{O}$ , 19.8% (two methoxys); mol. wt., 313.

*Melicopine* gave a positive methylenedioxy test (green→blue) when warmed with gallic acid in concentrated sulphuric acid. The alkaloid isolated as described above was always yellow, the colour being unaltered by repeated crystallization, with or without charcoal. However, after treatment with nitrous acid or potassium permanganate in acetone(12), it was recovered as pale yellow needles which still melted at 178.5–179.5 °C. Evidently the material isolated directly from the plant contains a trace of coloured impurity sufficient to affect the colour without having any significant effect on the melting point. *Melicopine* is readily soluble in chloroform, but sparingly in cold alcohol. It is a weak base, only sparingly soluble in 1% hydrochloric acid at room temperature, though it dissolves appreciably on heating and the free base crystallizes on cooling. It dissolved readily in 10% hydrochloric acid at room temperature but when the acid concentration was reduced by dilution to 5%, much of the alkaloid separated. Addition of ethereal hydrochloric acid to a solution of the base in ether caused precipitation of a hydrochloride which was easily hydrolysed to the base and underwent decomposition on heating. A picrate could not be prepared either in methanol or benzene solution, nor was a picrolonate formed in methanol.

Mother liquors from the crystallization of *melicopine* were combined with those from *melicopidine* and *melicopicine* for further treatment.

(c) *Ether-Soluble Fraction*

The combined chloroform-ether filtrates after precipitation of *melicopine* were treated with ethereal sulphuric acid until no more precipitate was formed. The ether was decanted and the sticky red precipitate washed by decantation with ether. The decantate was washed with water, distilled, and the residue set aside for purification of the triterpene. The alkaloid salts were decomposed with water, the mixture basified, and extracted with chloroform. The chloroform was removed and the sticky residue taken up in the minimum amount of methanol. On standing, yellow crystals separated. These were filtered, washed with cold methanol, and the filtrates concentrated and allowed to crystallize. The solid was again filtered, washed, and set aside. This was repeated until as much clean crystalline material as possible had been separated, when the filtrate was combined with the mother liquors from the crystallization of *melicopine* and set aside for further treatment.

The yellow crystalline solid was finely ground and triturated with 10% hydrochloric acid (60–75 ml./10 g.). This dissolved the bulk of the mixed alkaloids, the residue being added to previous mother liquors. The acid concentration was then reduced to 5% by dilution with an equal volume of water, resulting in the almost complete separation of *melicopicine*, which was filtered off after about an hour. The filtrate was then further diluted until the acid concentration was about 2% when long orange needles of *melicopidine hydrochloride* crystallized. The solution was filtered, and the residual alkaloid, recovered by basification and extraction with chloroform,

was combined with the mother liquors from the crystallization of melicopidine and melicopine and further quantities of these two alkaloids obtained by repeating the separation with acid.

#### (d) *Melicopidine*

The separated hydrochloride was decomposed with water, basified, and the alkaloid taken into chloroform. Removal of the solvent and crystallization of the residue from methanol afforded melicopidine as pale yellow prisms, m.p. 121–122 °C.

Found: C, 65.1; H, 4.9; N, 4.5;  $\text{CH}_3\text{N}$ , 8.5;  $\text{CH}_3\text{O}$ , 19.5%; mol. wt. (11), 315;  $[\alpha]_{\text{D}}$ , 0° in chloroform.

Calculated for  $\text{C}_{17}\text{H}_{15}\text{O}_5\text{N}$ : C, 65.2; H, 4.8; N, 4.5;  $\text{CH}_3\text{N}$ , 9.3;  $\text{CH}_3\text{O}$ , 19.8% (two methoxyls); mol. wt., 313.

Melicopidine gave a positive test for the methylenedioxy group on warming with gallic acid in sulphuric acid. It dissolved readily in organic solvents other than light petroleum. Though sparingly soluble in cold 1% hydrochloric acid it dissolved easily on heating and the solution when cooled deposited long orange needles of melicopidine hydrochloride, m.p. 88–90 °C. (with effervescence). The hydrochloride could also be precipitated from ethereal solution. It was hydrolysed by water. A sample dried under reduced pressure at room temperature yielded 90.0% melicopidine,  $\text{C}_{17}\text{H}_{15}\text{O}_5\text{N} \cdot \text{HCl}$  requires 89.6%. The hydrochloride, though not very soluble in 1.4% aqueous hydrochloric acid, dissolves readily at higher acid concentrations. On mixing methanolic solutions of melicopidine and picric acid *melicopidine picrate* crystallized. It was obtained as orange needles after recrystallization from methanol, m.p. 134–135 °C.

Found: liberated base, 58.3; C, 51.1; H, 3.2; N, 10.3;  $\text{CH}_3\text{O}$ , 11.5%.

Calculated for  $\text{C}_{17}\text{H}_{15}\text{O}_5\text{N} \cdot \text{C}_6\text{H}_3\text{O}_7\text{N}_3$ :  $\text{C}_{17}\text{H}_{15}\text{O}_5\text{N}$ , 57.7; C, 50.9; H, 3.3; N, 10.3;  $\text{CH}_3\text{O}$ , 11.4%.

The *picrolonate* crystallized from methanol as orange needles, m.p. 153–154 °C.

Found: C, 56.2; H, 3.7; N, 11.8%.

Calculated for  $\text{C}_{17}\text{H}_{15}\text{O}_5\text{N} \cdot \text{C}_{10}\text{H}_8\text{O}_6\text{N}_4$ : C, 56.2; H, 4.0; N, 12.1%.

#### (e) *Melicopine*

Melicopine crystallized from methanol as yellow prisms, m.p. 133–134 °C.

Found: C, 65.7; H, 5.8; N, 4.5;  $\text{CH}_3\text{N}$ , 9.3;  $\text{CH}_3\text{O}$ , 37.7%;  $[\alpha]_{\text{D}}$ , 0° in chloroform.

Calculated for  $\text{C}_{18}\text{H}_{19}\text{O}_5\text{N}$ : C, 65.7; H, 5.8; N, 4.3;  $\text{CH}_3\text{N}$ , 8.8;  $\text{CH}_3\text{O}$ , 37.7% (four methoxyls).

Melicopine did not give a methylenedioxy test on warming with gallic acid in sulphuric acid. The solubility of melicopine in organic solvents resembled that of melicopidine, though melicopine is somewhat more soluble in light petroleum. It was only sparingly soluble in 1% hydrochloric acid at room temperature but dissolved on heating and, as with melicopine, the free base crystallized on cooling. Dilution of a solution in 10% hydrochloric acid resulted in the almost complete separation of melicopine when the acid concentration reached 5%, the separation at this point being much greater than with melicopine. Melicopine did not form a *picrate* or a *picrolonate* in methanol, nor a *picrate* in chloroform, but the hydrochloride could be precipitated from ethereal solution. It was immediately hydrolysed by water.

#### (f) *Examination of Mother Liquors*

All alkaloid mother liquors were combined, the solvents removed, and the residue dissolved in benzene. The benzene solution was shaken successively (2–3 times each) with 1, 2½, 5, 10, and 20% hydrochloric acid. The extracts were basified and examined separately. The 20% hydrochloric acid extract contained melicopine, the 10% a mixture of melicopine and melicopine, the 5 and 2½% mixtures of melicopine and melicopidine. No other alkaloids could be detected in these fractions. The 1% hydrochloric acid extract was basified, the alkaloids extracted with chloroform, the solvent removed, and the residue again dissolved in benzene (smaller volume). The benzene solution was shaken with 1, 2½, and 5% hydrochloric acid. From the 2½ and 5% hydrochloric acid extracts were obtained some melicopidine and a little melicopine. The 1%

hydrochloric acid extract yielded a colourless base, *acronycidine*, which crystallized from methanol as needles or large prisms, m.p. 136.5–137.5 °C.

Found: C, 62.4; H, 5.2; N, 4.9%.

Calculated for  $C_{15}H_{16}O_5N$ : C, 62.3; H, 5.2; N, 4.8%.

There was also obtained from the 1% hydrochloric acid extract a small amount (c. 40 mg. from 22 kg. bark) of an impure almost colourless base which melted over the range 198–205 °C. (decomp., uncorr.) and which was not further investigated.

(g) *Triterpene Fraction—Identification of Lupeol*

After distillation of the ether from which the mixed alkaloid salts had been precipitated, the residue could be crystallized directly, but it was better purified by first refluxing with alcoholic potash. The recovered triterpene melted at 215–216 °C. after several crystallizations from ethyl acetate from which it separated as colourless needles. The mixed m.p. with an authentic specimen of lupeol was 214–216 °C.

Found: C, 84.6; H, 11.6%.

Calculated for  $C_{30}H_{50}O$ : C, 84.5; H, 11.7%.

$[\alpha]_D^{18}$ , +27.4° (c 1.17 in chloroform), whence  $[M]_D$ , +116.7°. The substance gave a violet-red colour changing to brown when a solution in acetic anhydride was treated with concentrated sulphuric acid. Acetylation gave *lupeol acetate*, colourless needles from acetone, m.p. 218–218.5 °C., mixed m.p. with an authentic specimen 217–218 °C.

Found: C, 81.8; H, 10.7%.

Calculated for  $C_{30}H_{48}OOC.CH_3$ : C, 82.1; H, 11.1%.

$[\alpha]_D^{18}$ , +41.7° (c 1.224 in chloroform), whence  $[M]_D$ , +195.2°.

Benzoylation with benzoyl chloride in pyridine at room temperature gave *lupeol benzoate*, colourless needles from ethanol, m.p. 268–271 °C. (uncorr.).

Found: C, 83.7; H, 10.3%.

Calculated for  $C_{30}H_{48}OOC.C_6H_5$ : C, 83.8; H, 10.2%.

$[\alpha]_D^{18}$ , +59.3° (c 1.333 in chloroform), whence  $[M]_D$ , +314.3°. The values for the rotations of lupeol, its acetate and benzoate, and for the molecular rotation differences resulting from acetylation and benzoylation agree well with those recorded by Barton and Jones(13).

(h) *Determination of the Approximate Alkaloid Content of the Bark*

750 g. milled bark was extracted exhaustively with methanol (soxhlet). The methanol was distilled, the residue treated with water and chloroform, and the chloroform washed with caustic soda. After removal of the chloroform the alkaloids were extracted by boiling five times with 2% hydrochloric acid (c. 5 l. in all). Weakly basic alkaloids were removed by chloroform from the acid solution which was then basified and again extracted with chloroform. The latter extract yielded 0.5 g. crude acronycidine of which a further 0.3 g. was separated from the weakly basic alkaloid fraction. Hence the acronycidine content was approximately 0.1%. The weakly basic alkaloids totalled 22.5 g., i.e. 3.0%. Separation along the lines of the isolation procedure described above gave the following amounts of crude alkaloids:

melicopine	6.5 g.	(0.9%)
melicopidine	7.1 g.	(0.9%)
melicopicine	6.8 g.	(0.9%)

There remained 1.9 g. of mixed bases. Evidently this specimen of bark contained approximately 1% of each of the three main alkaloids and 0.1% acronycidine. The acid-insoluble material, largely lupeol, weighed 8.4 g., equivalent to 1.1%. These figures represent an average specimen of the bark examined, but the total alkaloid content and the proportions of the alkaloids varied somewhat. A specimen of root bark contained 4.3% total alkaloids. From this was isolated melicopine 0.9%, melicopidine 2.0%, melicopicine 0.6%, and acronycidine 0.25%.

(i) *Isolation of Alkaloids from the Leaves*

Milled leaves were exhaustively extracted with methanol, the solvent removed, and the viscous residue extracted six–eight times with boiling 2% hydrochloric acid. The hydrochloric



acid solutions were shaken with chloroform, then basified, and again extracted with chloroform. The following were obtained from the weakly basic fraction :

(i) Melicopine (c. 0.2%) m.p. and mixed m.p. with a specimen from the bark 178–179 °C.  
Found : C, 65.3; H, 5.0; N, 4.6%.

Calculated for  $C_{17}H_{15}O_5N$  : C, 65.2; H, 4.8; N, 4.5%.

(ii) Melicopidine (c. 0.8%) m.p. and mixed m.p. 121–122 °C.

(iii) Melicopieine (c. 0.03%) m.p. and mixed m.p. 132–133 °C.

(iv) Almost colourless material readily soluble in 1% hydrochloric acid. This was combined with the more strongly basic fraction and the whole recrystallized from methanol (charcoal), giving skimmianine 0.3%. Mother liquors from all fractions from which further crystalline material was not readily separable, represented approximately 0.2%. Thus the total alkaloid content was about 1.5%. The skimmianine was obtained as colourless prisms, m.p. 178–179 °C., not depressed by admixture with an authentic specimen.

Found : C, 64.7; H, 5.2;  $CH_3O$ , 35.3%.

Calculated for  $C_{14}H_{13}O_4N$  : C, 64.9; H, 5.0;  $CH_3O$ , 35.9% (three methoxyls).

The *picrate* crystallized from methanol as glistening yellow needles, m.p. 197–198 °C.

Found : C, 49.3; H, 3.2; N, 11.5%.

Calculated for  $C_{14}H_{13}O_4N.C_6H_5O_7N_3$  : C, 49.2; H, 3.3; N, 11.5%.

The base was heated with methyl iodide in a sealed tube at 100 °C. for four hours. The product crystallized from water as colourless needles, m.p. 190–191 °C. Deulofeu, Labriola, and de Langhe(10) record the following melting points (not stated to be corrected) for skimmianine from *Fagara coco* : base, 178 °C. : *picrate*, 195 °C. : *isoskimmianine*, 188–189 °C. From another sample of leaves a little acronycidine was obtained from the mother liquors after crystallization of the skimmianine. It was purified by recrystallization of the *picrate*, from which the base was recovered and obtained as colourless needles, m.p. and mixed m.p. 135–137 °C.

### III. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Industrial Chemistry, C.S.I.R.O. The author wishes to thank L. J. Webb, Division of Plant Industry, C.S.I.R.O., for supplying the material used in this investigation and for information concerning its origin : H. F. Haynes for assistance with the isolation of the alkaloids : R. D. Brown for carrying out molecular weight determinations : and Dr. V. Deulofeu for kindly supplying a specimen of skimmianine.

### IV. REFERENCES

- (1) BARNARD, C., and FINNEMORE, H.—*J. Coun. Sci. Industr. Res. Aust.* **18** : 277 (1945). See also *Coun. Sci. Industr. Res. Aust. Ann. Rep.* 1940–41, p. 14; 1941–42, p. 11; 1942–43, p. 11; 1944–45, p. 15; 1945–46, p. 15.
- (2) WEBB, L. J.—*Coun. Sci. Industr. Res. Aust. Bull. No.* 232 (1948).
- (3) HUGHES, G. K., LAHEY, F. N., PRICE, J. R., and WEBB, L. J.—*Nature* **162** : 223 (1948).
- (4) HENRY, T. A. "The Plant Alkaloids." 3rd Ed. (Blakiston Book Co. : Philadelphia, 1939.)
- (5) ENGELER, A., and PRANTL, K.—"Die Natürlichen Pflanzenfamilien." Vol. 19a. (Wilhelm Engelmann : Leipzig, 1931.)
- (6) LAHEY, F. N., and THOMAS, W. C.—*Aust. J. Sci. Res. A* **2** (in press).
- (7) JONES, T. G. H., and WHITE, M.—*Proc. Roy. Soc. Qd.* **41** : 154 (1929).
- (8) GOODSON, J. A.—*Biochem. J.* **15** : 123 (1921).
- (9) MOOKERJEE, A., and BOSE, P. K.—*J. Ind. Chem. Soc.* **23** : 1 (1946).
- (10) DEULOFEU, V., LABRIOLA, R., and DE LANGHE, J.—*J. Amer. Chem. Soc.* **64** : 2326 (1942).
- (11) CLARK, E. P.—*Industr. Engng. Chem. (Anal. Ed.)* **13** : 820 (1941).
- (12) CROW, W. D., and PRICE, J. R.—*Aust. J. Sci. Res. A* **2** : 255 (1949).
- (13) BARTON, D. H. R., and JONES, E. R. H.—*J. Chem. Soc.* **1944** : 659 (1944).



# ALKALOIDS OF THE AUSTRALIAN RUTACEAE: *MELICOPE FAREANA*

## II. PRELIMINARY EXAMINATION OF MELICOPINE, MELICOPIDINE, AND MELICOPICINE

By W. D. CROW\* and J. R. PRICE\*

[Manuscript received March 9, 1949]

### Summary

Four of the five oxygen atoms in the alkaloids melicopine, melicopidine, and melicopicine are present as alkoxy groups, either methoxy or methylenedioxy. In each alkaloid one of the methoxy groups is remarkably sensitive to hydrolysis by acids. The resulting hydroxy compound is an alcohol, or more likely, an unreactive phenol. The methylenedioxy group in melicopine and melicopidine is attacked by alcoholic alkali and replaced by an hydroxyl and an alkoxy group. The resulting phenols still contain the easily hydrolysed methoxy group, the loss of which gives rise to a series of dihydroxy compounds. A number of acidic oxidation products is described.

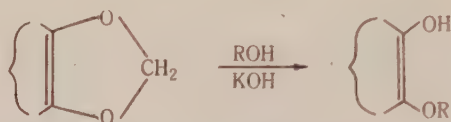
### I. INTRODUCTION

The isolation of the alkaloids melicopine, melicopidine, and melicopicine from the bark and leaves of *Melicope fareana* F. Muell. has been described in Part I(1). Melicopine and melicopidine were shown to be isomeric ( $C_{17}H_{15}O_5N$ ), each containing two methoxy groups, a methylenedioxy, and a methylimino group. Melicopicine ( $C_{18}H_{19}O_5N$ ) contains four methoxy groups and a methylimino group. All three alkaloids are recovered unchanged after attempted acetylation and after treatment with potassium permanganate in acetone. They fail to react with 2,4-dinitrophenylhydrazine, to yield quaternary salts with alkyl halides, and are unattacked by one mole of nitrous acid. They are insoluble in and are unattacked by prolonged boiling with aqueous alkali.

However, when melicopine or melicopidine is refluxed with ethanolic potash, each is converted to a yellow *phenol*. The phenols from the two alkaloids are isomeric, of empirical formula  $C_{18}H_{19}O_5N$ . They contain three alkoxy groups and no longer give positive methylenedioxy tests. Both are monohydric, forming *monobenzoates* and, in the case of the melicopidine derivative, a *monoacetate*. With methanolic potash melicopine and melicopidine again give isomeric trialkoxyphenols but with empirical formula  $C_{17}H_{17}O_5N$ . Melicopidine with *isopropanolic* potash gives a trialkoxyphenol,  $C_{19}H_{21}O_5N$ . The formulae of these phenols show that the solvent alcohol takes part in the reaction. Evidently the methylenedioxy ring opens according to the scheme shown below to give

\* Division of Industrial Chemistry, C.S.I.R.O.

hydroxyalkoxy- compounds—a reaction first reported by Robinson and Robinson(2) for 4-nitro-1,2-methylenedioxybenzene.



This being so, the products from methanolic potash are trimethoxyphenols and those from ethanolic potash ethoxydimethoxyphenols. These phenols are also formed, though much more slowly, when aqueous alcoholic alkali is employed.

The instability of the salts of the three alkaloids with mineral acids, has been referred to in Part I(1). These salts decompose slowly at room temperature or rapidly on heating to give mixtures of the original alkaloid with more highly coloured substances which are even weaker bases than the alkaloids. The same substances may be obtained in quantitative yield by refluxing the alkaloids with alcoholic hydrochloric acid. Analysis shows in each case a loss of  $\text{CH}_2$  as methoxyl, and the substances will be referred to as the *noralkaloids*. Melicopine gives *normelicopine* (red,  $\text{C}_{16}\text{H}_{13}\text{O}_5\text{N}$ ), melicopidine gives *normelicopidine* (orange-red,  $\text{C}_{16}\text{H}_{13}\text{O}_5\text{N}$ ), and melicopicine gives *normelicopicine* (orange,  $\text{C}_{17}\text{H}_{17}\text{O}_5\text{N}$ ), the first two still containing a methylenedioxy group. Acetylation of the *noralkaloids* gives pale yellow monoacetyl derivatives resembling the alkaloids in solubility, and methylation with dimethyl sulphate regenerates the alkaloids. The *noralkaloids* are insoluble in aqueous alkali, but addition of alkali to their alcoholic solutions causes a small but definite enhancement of colour. This behaviour, which is not shown by the alkaloids themselves, suggests an unreactive phenolic hydroxyl, though the presence of an alcoholic group cannot be excluded. They are extremely weak bases, soluble in concentrated hydrochloric acid, but precipitated on slight dilution. Unlike the alkaloids, the *norecompounds* are rapidly oxidized by permanganate in acetone to give a mixture of acids. They also differ from the alkaloids in that they are readily attacked by one mole of nitrous acid. The products are not nitroso compounds and the reaction will be discussed in a later paper. When the phenols obtained from melicopine and melicopidine with alcoholic potash are treated with alcoholic acid, one methoxyl group is demethylated, as in the parent alkaloids, giving rise to a series of *dihydroxy* compounds. These are orange-red in colour and each forms two acetyl derivatives, a *monoacetate* with acetic anhydride alone and a *diacetate* with acetic anhydride and anhydrous sodium acetate.

The ethoxydimethoxyphenols from melicopine and melicopidine are readily oxidized by hydrogen peroxide. That from melicopine gives an orange *acid*,  $\text{C}_{13}\text{H}_9\text{O}_4\text{N}$ , containing no alkoxyl group. The same acid is also obtained (a) as a by-product in the preparation of the dimethoxyethoxyphenol from melicopine, (b) from the corresponding dihydroxy compound from melicopine by oxidation of an alkaline solution with oxygen, and (c) from *normelicopine* and *normelicopidine* by the action of ethanolic potash, presumably due to fission of the methylenedioxy ring followed by atmospheric oxidation of the resulting

dihydroxy compounds. Oxidation of the ethoxydimethoxyphenol from melicopidine by hydrogen peroxide gives a colourless dibasic *acid*,  $C_{17}H_{17}O_7N$ , containing two alkoxyl groups. This acid is also formed as a by-product in the action of ethanolic potash on melicopidine. When the acid is heated at its melting point one carboxyl group is eliminated, giving an *acid*,  $C_{16}H_{17}O_5N$ . Further degradation also takes place to an *acid*,  $C_{13}H_{13}O_4N$ , which is found as a by-product in the action of aqueous ethanolic potash on melicopidine.

The dihydroxy compounds obtained from melicopidine by treatment with methanolic, ethanolic, or isopropanolic potash, followed by alcoholic acid, all give the same product when oxidized in alkaline solution by air or oxygen—a colourless dibasic *acid*,  $C_{13}H_9O_6N$ . Other minor acid by-products which were encountered in the alcoholic potash reactions and in the oxidation experiments are described in the experimental section and will be referred to in later papers.

## II. EXPERIMENTAL

All melting points are corrected unless otherwise stated. Microanalyses were carried out by R. B. Bradbury and D. J. Clark.

### (a) General Properties of the Alkaloids

All three alkaloids were insoluble in aqueous alkalis. They were recovered unchanged (m.p. and mixed m.p.) after: (i) refluxing with 10% caustic soda for four hours, (ii) attempted acetylation by refluxing with acetic anhydride and either pyridine or anhydrous sodium acetate, (iii) attempted preparation of quaternary salts with methyl and ethyl iodides at room temperature or under reflux, (iv) treatment of hydrochloric acid solutions with one mole of nitrous acid at 0 °C., (v) treatment with alcoholic 2,4-dinitrophenylhydrazine hydrochloride. Melicopidine, however, apparently formed an addition compound and was recovered only after decomposing the 2,4-dinitrophenylhydrazine with cold alkali, (vi) treatment with potassium permanganate in boiling acetone for 20 hours.

### (b) Reactions with Ethanolic Potash

(i) *Melicopine*.—The alkaloid (5 g.) in ethanol (250 ml.) containing potassium hydroxide (10 g.) was refluxed for 10 hours, the colour of the solution changing to deep orange-red. The bulk of the solvent was removed, the residue diluted with water and extracted repeatedly with ether. The combined ethereal extracts were washed with 0.5% aqueous potash and water, the washings being added to the alkaline solution. The yellow ethereal extract was dried and evaporated and the small amount of solid alkali-insoluble residue reserved for future examination. The aqueous alkaline solution was saturated with carbon dioxide, precipitating a yellow crystalline solid (3.1 g.). A little more of this phenol was obtained by extraction of the filtrate with chloroform. The aqueous bicarbonate was set aside for examination of acid by-products (*vide infra*).

The phenol crystallized from methanol as yellow prisms, m.p. 147–149 °C.

Found: C, 65.7; H, 5.8; N, 4.3;  $CH_3O$ , 27.7%.

Calculated for  $C_{18}H_{19}O_5N$ : C, 65.7; H, 5.8; N, 4.3;  $CH_3O$ , 28.3% (two methoxyls and one ethoxyl calculated as methoxyl).

The phenol gave a brownish-red colour with ferric chloride. The methylenedioxy test was negative. Benzoylation (Schotten-Baumann) gave the *benzoate*, pale yellow needles from ethanol, m.p. 182–184 °C., which darken on exposure to light.

Found: C, 69.3; H, 5.1; N, 3.2;  $CH_3O$ , 20.9%.

Calculated for  $C_{18}H_{18}O_4N(OOC.C_6H_5)$ : C, 69.3; H, 5.3; N, 3.2;  $CH_3O$ , 21.5% (two methoxyls and one ethoxyl calculated as methoxyl).



The phenol was recovered unchanged (m.p. and mixed m.p.) after attempted preparation of a *p*-nitrophenylhydrazone. Methylation with dimethyl sulphate and caustic soda gave a non-crystallizable gum.

The aqueous bicarbonate solution after removal of the phenol was acidified with hydrochloric acid and allowed to stand. The acid which separated was obtained as orange-yellow needles, m.p. 262–263 °C. (uncorr.) after crystallization from acetic acid.

Found: C, 64.0; H, 4.1; N, 5.9%.

Calculated for  $C_{13}H_9O_4N$ : C, 64.2; H, 3.7; N, 5.8%.

The filtrate from this acid was extracted repeatedly with chloroform and the chloroform extracts distilled leaving a dark semi-solid residue. Washing with benzene and then methanol gave a yellow solid which was obtained as straw coloured needles after crystallization from acetic acid. The substance melted at 291–293 °C. (uncorr.) after sintering at 280 °C.

(ii) *Melicopidine*.—The same procedure was employed as with melicopine. The reaction appeared to proceed more rapidly and the yield of phenol was higher (75–90%). It crystallized from alcohol in glistening yellow rectangular plates, m.p. 181.5–182.5 °C.

Found: C, 65.5; H, 5.6; N, 4.2;  $CH_3O$ , 27.9%.

Calculated for  $C_{18}H_{19}O_5N$ : C, 65.7; H, 5.8; N, 4.3;  $CH_3O$ , 28.3% (two methoxyls and one ethoxyl calculated as methoxyl).

The phenol did not give a methylenedioxy test but gave a brownish-red colour with ferric chloride. The benzoate crystallized from aqueous ethanol as bright yellow, stout needles, m.p. 132–133 °C.

Found: C, 69.5; H, 5.3; N, 3.3;  $CH_3O$ , 20.8%.

Calculated for  $C_{18}H_{18}O_4N(OOC.C_6H_5)$ : C, 69.3; H, 5.3; N, 3.2;  $CH_3O$ , 21.5% (two methoxyls and one ethoxyl calculated as methoxyl).

Acetylation of the phenol with acetic anhydride and pyridine gave a noncrystallizable gum, but a satisfactory product was obtained when the phenol (0.5 g.) dissolved in acetic anhydride (3 ml.) containing concentrated sulphuric acid (3 drops) was allowed to stand at room temperature for two days. The acetate crystallized from ethyl acetate-light petroleum or aqueous ethanol as pale yellow needles, m.p. 129–131 °C.

Found: C, 64.6; H, 5.8; N, 4.0;  $CH_3CO$ , 12.0%.

Calculated for  $C_{18}H_{18}O_4N(OOC.CH_3)$ : C, 64.7; H, 5.7; N, 3.8;  $CH_3CO$ , 11.6%.

Hydrolysis of the acetate by aqueous methanolic potash (24%) at room temperature regenerated the phenol (identified by m.p. and mixed m.p.). The phenol was recovered unchanged (m.p. and mixed m.p.) after attempted preparation of an oxime.

In the preparation of the phenol a little unchanged melicopidine (m.p. and mixed m.p.) was found in the alkali insoluble fraction, together with *normelicopidine* (*vide infra*), reddish-orange needles from chloroform alcohol, m.p. 209–210 °C., mixed m.p. 210–211 °C. Acidification of the aqueous bicarbonate solution after removal of the phenol gave a colourless crystalline acid. More of this was obtained by extraction of the filtrate with chloroform. Recrystallization (charcoal) from water or aqueous methanol gave colourless needles, m.p. 198–199 °C. (decomp.).\*

Found: C, 58.8; H, 5.0%.

Calculated for  $C_{17}H_{17}O_7N$ : C, 58.8; H, 4.9%.

The reaction with melicopidine was also carried out with aqueous ethanolic potash. The alkaloid (1 g.) was refluxed for 40 hours with a mixture of water (100 ml.) and ethanol (30 ml.) containing potassium hydroxide (10 g.). The bulk of the melicopidine was recovered unchanged (m.p. and mixed m.p.), but a small amount (c. 5%) of phenol was formed. This was the same (m.p. and mixed m.p.) as that obtained with ethanolic potash. The acid fraction in this experiment crystallized from ethyl acetate as colourless leaflets which sintered at 191 °C. and melted at 195–197 °C. The mixed m.p. with the acid,  $C_{17}H_{17}O_7N$ , obtained as a by-product in the alcoholic

\* This temperature varies with the rate of heating. The figure recorded was obtained by placing the m.p. tube in a bath preheated to 175–180 °C.



potash reaction was depressed c. 20 °C., but there was no depression on admixture with the acid,  $C_{18}H_{13}O_4N$ , described below.

(c) *Reactions with Methanolic Potash*

(i) *Melicopine*.—The alkaloid (2 g.) was refluxed for 50 hours with methanol (125 ml.) containing potassium hydroxide (8 g.). Reaction occurred very slowly, as shown by the slow colour change, and the non-phenolic fraction contained much of the starting material (m.p. and mixed m.p.). The yield of phenol varied considerably (5–40%). It was obtained as small bright yellow prisms from methanol, m.p. 190.5–191.5 °C.

Found: C, 65.1; H, 5.6; N, 4.4;  $CH_3O$ , 28.9%.

Calculated for  $C_{17}H_{17}O_5N$ : C, 64.8; H, 5.4; N, 4.4;  $CH_3O$ , 29.5% (three methoxyls).

The acidic by-product, orange-yellow needles from acetic acid, melted at 262–263 °C. (uncorr.) alone or mixed with the acid,  $C_{13}H_9O_4N$ , obtained from melicopine and ethanolic potash.

(ii) *Melicopidine*.—The substance (8 g.) was refluxed for 20 hours with methanol (250 ml.) containing potassium hydroxide (20 g.). 2.5 g. starting material was recovered (m.p. and mixed m.p.). The phenol (4.5 g., 56%), yellow needles from aqueous methanol, melted at 165–166 °C.

Found: C, 64.7; H, 5.5; N, 4.5;  $CH_3O$ , 28.9%.

Calculated for  $C_{17}H_{17}O_5N$ : C, 64.8; H, 5.4; N, 4.4;  $CH_3O$ , 29.5% (three methoxyls).

The phenol did not give a methylenedioxy test. The small acidic fraction was not investigated. A small yield (c. 5%) of the same phenol (m.p. and mixed m.p.) was obtained using aqueous methanolic potash. Melicopidine (0.5 g.) was refluxed for 36 hours with a mixture of methanol (50 ml.) and water (150 ml.) containing potassium hydroxide (15 g.). The bulk of the alkaloid was recovered unchanged (m.p. and mixed m.p.).

(d) *The Action of isoPropanolic Potash on Melicopidine\**

Melicopidine (2 g.) was refluxed for one hour with *isopropanol* (100 ml.) containing potassium hydroxide (5 g.). The reaction, gauged by the change in colour, took place much faster than with the lower alcohols, and no unchanged alkaloid was recovered. The phenol, obtained in excellent yield, crystallized from ethanol in small bright yellow prisms, m.p. 176–178 °C.

Found: C, 66.7; H, 6.2; N, 4.1;  $CH_3N$ , 7.9;  $CH_3O$ , 27.6%.

Calculated for  $C_{19}H_{21}O_5N$ : C, 66.5; H, 6.1; N, 4.1;  $CH_3N$ , 8.5;  $CH_3O$ , 27.1% (two methoxyls and one *isopropoxyl* calculated as methoxyl).

The acid fraction was not investigated.

(e) *Reactions with Alcoholic Hydrochloric Acid*

(i) *Melicopine*.—The alkaloid (10 g.) was refluxed with a mixture of ethanol (500 ml.) and concentrated hydrochloric acid (50 ml.). Red crystals soon appeared. After one hour these were filtered and the mother liquor concentrated giving a further quantity of the same material. The substance was not a hydrochloride as it was unaffected by aqueous alkali. Recrystallization from chloroform-alcohol gave *normelicopine*, red needles m.p. 235.5–236.5 °C., yield 96%.

Found: C, 64.2; H, 4.4; N, 4.9;  $CH_3O$ , 10.4%.

Calculated for  $C_{16}H_{13}O_5N$ : C, 64.2; H, 4.3; N, 4.7;  $CH_3O$ , 10.4% (one methoxyl).

The substance gave a positive methylenedioxy test. Acetylation with boiling acetic anhydride and either pyridine or anhydrous sodium acetate gave the *acetate*, yellow needles from chloroform-ethanol, m.p. 256–258 °C.

Found: C, 63.6; H, 4.5; N, 4.2;  $CH_3CO$ , 12.7%.

Calculated for  $C_{16}H_{12}O_4N(OOC.CH_3)$ : C, 63.3; H, 4.4; N, 4.1;  $CH_3CO$ , 12.6%.

*norMelicopine* dissolved readily in chloroform, sparingly in alcohol, and was insoluble in water. It dissolved in concentrated hydrochloric acid giving an intense red solution which deposited deep red crystals of a hydrochloride almost immediately. This decomposed on dilution with water, *normelicopine* being almost insoluble in 25% hydrochloric acid. *norMelicopine* was

insoluble in aqueous alkali, but addition of alcoholic potash to an alcoholic solution caused a marked deepening in colour. *nor*Melicopine (4 g.) was refluxed for 32 hours with ethanol (250 ml.) containing potassium hydroxide (20 g.). Some unchanged starting material (0.9 g., m.p. and mixed m.p.) was recovered. No phenol was obtained, but the bicarbonate solution on acidification and heating gave a crystalline precipitate of an orange acid (1.4 g.), m.p. 261–262 °C. (uncorr.), mixed m.p. with the acid,  $C_{13}H_9O_4N$ , 262–263 °C. (uncorr.). Extraction of the filtrate with chloroform and repeated crystallization of the extracted acid gave a small amount of a substance which sintered at 280 °C. and melted at 294–296 °C. (uncorr.). *nor*Melicopine was recovered unchanged (m.p. and mixed m.p.) after treatment with *o*-phenylenediamine in acetic acid. It was oxidized by permanganate in acetone solution, but the product was a mixture of acids which could not be separated. *nor*Melicopine was methylated by refluxing 1 g. in acetone (150 ml.) with dimethyl sulphate (two portions: 10 ml., 5 ml.) and excess anhydrous potassium carbonate for 24 hours. The yellow precipitate after removal of acetone and addition of water was dissolved in benzene and the benzene solution washed with 1% hydrochloric acid (discarded) then extracted several times with 10% hydrochloric acid. Unchanged *normelicopine* (0.27 g.) was recovered from the benzene, and the 10% acid extracts after basification yielded melicopine (0.64 g.), m.p. and mixed m.p. after crystallization from ethanol 178–179 °C.

Found: C, 65.3; H, 4.9; N, 4.5;  $CH_3O$ , 19.7%.

Calculated for  $C_{17}H_{15}O_5N$ : C, 65.2; H, 4.8; N, 4.5;  $CH_3O$ , 19.8%.

(ii) *Melicopidine*.—*nor*Melicopidine was prepared by the method described for *normelicopine*, the yield being 93%. It crystallized from chloroform-ethanol in orange-red flat needles, m.p. 211–212 °C.

Found: C, 64.4; H, 4.6; N, 4.7;  $CH_3O$ , 10.6%.

Calculated for  $C_{16}H_{13}O_5N$ : C, 64.2; H, 4.3; N, 4.7;  $CH_3O$ , 10.4% (one methoxyl).

The *acetate*, prepared in the usual manner, crystallized from ethyl acetate-light petroleum as yellow plates, m.p. 183–184 °C.

Found: C, 63.5; H, 4.6; N, 4.2;  $CH_3CO$ , 13.1%.

Calculated for  $C_{16}H_{13}O_4N(OOC.CH_3)$ : C, 63.3; H, 4.4; N, 4.1;  $CH_3CO$ , 12.6%.

*nor*Melicopidine (identified by m.p. and mixed m.p.) was also obtained in c. 50% yield by heating melicopidine hydrochloride in the dry state at 100 °C. for a few minutes. It resembled *normelicopine* in solubility. A solution in concentrated hydrochloric acid was orange-red and the hydrochloride was more soluble than that of *normelicopine*. The colour of an alcoholic solution of *normelicopidine* was intensified by addition of alcoholic potash. *nor*Melicopidine (4 g.) was refluxed for 60 hours with ethanol (250 ml.) containing caustic potash (20 g.). The products (identified by m.p. and mixed m.p.) were the same as from *normelicopine* but in different relative amounts:  $C_{13}H_9O_4N$ , 0.6 g.; acid, m.p. 294–296 °C., 0.6 g.; unchanged *normelicopidine*, 1.9 g. *nor*Melicopidine was oxidized by permanganate in acetone, but again the result was a mixture of acids which could not be separated satisfactorily.

Methylation of *normelicopidine* with dimethyl sulphate in acetone in the presence of anhydrous potassium carbonate gave melicopidine. The base was purified through the hydrochloride (twice) and identified by m.p. and mixed m.p. 121–122 °C.

Found: C, 65.4; H, 4.8; N, 4.6%.

Calculated for  $C_{17}H_{15}O_5N$ : C, 65.2; H, 4.8; N, 4.5%.

(iii) *Melicopicine*. *nor*Melicopicine separated from the alcoholic reaction mixture only after concentration and dilution with water. It crystallized from aqueous methanol as orange needles, m.p. 129–129.5 °C. Yield 92%.

Found: C, 64.8; H, 5.5; N, 4.4;  $CH_3O$ , 29.1%.

Calculated for  $C_{17}H_{17}O_5N$ : C, 64.8; H, 5.4; N, 4.4;  $CH_3O$ , 29.5%.

Acetylation could not be effected by acetic anhydride and pyridine, but the *acetate* was obtained by refluxing with acetic anhydride and anhydrous sodium acetate. It crystallized from chloroform-light petroleum as yellow prisms, m.p. 113–114.5 °C.

Found: C, 63.9; H, 5.4; N, 3.9;  $CH_3CO$ , 12.4%.

Calculated for  $C_{17}H_{16}O_4N(OOC.CH_3)$ : C, 63.9; H, 5.3; N, 3.9;  $CH_3CO$ , 12.0%.

*nor*Melicopine is more soluble in organic solvents than the other two *nor*alkaloids. The hydrochloride was orange and its solutions in concentrated hydrochloric acid orange-red. The colour of an alcoholic solution deepened on addition of alcoholic potash. *nor*Melicopine was easily oxidized by permanganate in acetone giving a mixture of acids which were not obtained pure. Methylation by shaking with caustic soda (20%) and excess dimethyl sulphate and finally warming on the water-bath gave melicopine, m.p. and mixed m.p. 133–134 °C., in good yield.

Found: C, 65.7; H, 5.8; N, 4.5%.

Calculated for  $C_{18}H_{19}O_5N$ : C, 65.7; H, 5.8; N, 4.3%.

(f) *Action of Hydrochloric Acid on the Alkoxydimethoxyphenols*

(i) The ethoxydimethoxyphenol obtained from melicopine and ethanolic potash was refluxed with 7% ethanolic hydrochloric acid for three hours. The bulk of the alcohol was evaporated, the mixture cooled, and the crystalline dihydroxy compound filtered off. The filtrate was diluted with water, extracted with chloroform, and the chloroform extract washed with aqueous sodium bicarbonate which removed a small amount of coloured acidic material. The chloroform was evaporated and the residue combined with the main product (yield 90–95%). The dihydroxy compound crystallized from chloroform-ethanol as orange-red leaflets, m.p. 201–202 °C.

Found: C, 65.2; H, 5.5; N, 4.5;  $CH_3O$ , 19.3%.

Calculated for  $C_{17}H_{17}O_5N$ : C, 64.8; H, 5.4; N, 4.4;  $CH_3O$ , 19.7% (one methoxyl and one ethoxyl calculated as methoxyl).

Acetylation with warm acetic anhydride gave the *monoacetate*, orange-yellow needles from ethanol, m.p. 155.5–156.5 °C.

Found: C, 63.8; H, 5.2;  $CH_3CO$ , 11.7%.

Calculated for  $C_{17}H_{16}O_4N(OOC.CH_3)$ : C, 63.9; H, 5.3;  $CH_3CO$ , 12.0%.

Further acetylation with acetic anhydride and anhydrous sodium acetate gave the *diacetate*, cream leaflets from ethyl acetate-light petroleum, m.p. 152–153 °C.

Found: C, 63.0; H, 5.2;  $CH_3CO$ , 22.0%.

Calculated for  $C_{17}H_{15}O_3N(OOC.CH_3)_2$ : C, 63.2; H, 5.3;  $CH_3CO$ , 21.6%.

(ii) The ethoxydimethoxyphenol from melicopidine and ethanolic potash, treated with alcoholic acid as above, gave a dihydroxy compound (yield 90–95%) and a small amount of a red acidic by-product. This dihydroxy compound is dimorphous. It separated from hot alcohol as red needles which on standing in contact with the solution at room temperature changed to lighter orange-red prisms. The crystals separating from a cooled solution were of this second type. Both forms melted at 172–173 °C.

Found: C, 64.8; H, 5.5; N, 4.4;  $CH_3O$ , 19.1%.

Calculated for  $C_{17}H_{17}O_5N$ : C, 64.8; H, 5.4; N, 4.4;  $CH_3O$ , 19.7% (one methoxyl and one ethoxyl calculated as methoxyl).

Acetic anhydride gave the *monoacetate*, orange-yellow needles from ethanol, m.p. 149–151 °C.

Found: C, 63.9; H, 5.4; N, 3.9;  $CH_3CO$ , 12.5%.

Calculated for  $C_{17}H_{16}O_4N(OOC.CH_3)$ : C, 63.9; H, 5.3; N, 3.9;  $CH_3CO$ , 12.0%.

Further acetylation with acetic anhydride and anhydrous sodium acetate gave the *diacetate*, cream leaflets from ethyl acetate-light petroleum, m.p. 172–173 °C.

Found: C, 62.9; H, 5.5; N, 3.5;  $CH_3CO$ , 21.7%.

Calculated for  $C_{17}H_{15}O_3N(OOC.CH_3)_2$ : C, 63.2; H, 5.3; N, 3.5;  $CH_3CO$ , 21.6%.

The acid by-product in the preparation of the dihydroxy compound crystallized from aqueous ethanol in masses of fine needles, m.p. 197–198 °C.

Found: C, 64.3; H, 4.6; N, 4.7%.

Calculated for  $C_{16}H_{13}O_5N$ : C, 64.2; H, 4.3; N, 4.7%.

(iii) The trimethoxyphenol from melicopine and methanolic potash on treatment with alcoholic hydrochloric acid gave a dimethoxydihydroxy compound which crystallized from chloroform-alcohol as flat orange-red needles, m.p. 242–243 °C.



Found: C, 64.0; H, 5.2; N, 4.6;  $\text{CH}_3\text{O}$ , 20.8%.

Calculated for  $\text{C}_{16}\text{H}_{15}\text{O}_5\text{N}$ : C, 63.8; H, 5.0; N, 4.7;  $\text{CH}_3\text{O}$ , 20.6% (two methoxyls).

(iv) The trimethoxyphenol from melicopidine and methanolic potash gave a dimethoxy-dihydroxy compound which separated as orange-red needles from methanol, m.p. 162–163 °C.

Found: C, 64.0; H, 5.0%.

Calculated for  $\text{C}_{16}\text{H}_{15}\text{O}_5\text{N}$ : C, 63.8; H, 5.0%.

(v) The isopropoxydimethoxyphenol from melicopidine and isopropanolic potash gave, on treatment with alcoholic acid, a dihydroxy compound still containing the isopropoxyl group. It crystallized from ethanol as red prisms, m.p. 150–151 °C.

Found: C, 65.6; H, 5.9; N, 4.5%.

Calculated for  $\text{C}_{18}\text{H}_{19}\text{O}_5\text{N}$ : C, 65.7; H, 5.8; N, 4.3%.

#### (g) Oxidation of the Ethoxydimethoxyphenols

(i) The ethoxydimethoxyphenol obtained from melicopine and ethanolic potash was oxidized in boiling 5% caustic potash solution by hydrogen peroxide (3%), added dropwise until the initial orange colour vanished. The solution was cooled and acidified, but no solid separated. On heating the acid solution the colour deepened and a voluminous crystalline precipitate formed. Recrystallization from acetic acid gave orange needles, m.p. 262–263 °C. (uncorr.), not depressed by admixture with the acid by-product of the same m.p. from the action of ethanolic potash on melicopine. Yield 50–60%.

Found: C, 64.3; H, 3.9; N, 5.9%.

Calculated for  $\text{C}_{13}\text{H}_9\text{O}_4\text{N}$ : C, 64.2; H, 3.7; N, 5.8%.

The substance contains no methoxyl group and is a weak acid, slowly extracted by chloroform from a bicarbonate solution. Solutions in alkali or in chloroform are yellow with a strong green fluorescence.

(ii) Oxidation of the ethoxydimethoxyphenol from melicopidine was carried out with hydrogen peroxide, as described above. Acidification of the cooled alkaline solution led to the immediate separation of a colourless acid, more of which was obtained by extraction of the mother liquors with chloroform. Recrystallization from water or aqueous methanol gave colourless needles, m.p. 198–199 °C. (decomp.).\* The substance is identical with the by-product from the action of ethanolic potash on melicopidine; a mixture of the two behaved in the same manner on heating. Yield 70–80%.

Found: C, 59.0; H, 4.9; N, 4.2;  $\text{CH}_3\text{O}$ , 17.9%; eq. wt. (by titration), 181.

Calculated for  $\text{C}_{11}\text{H}_7\text{O}_4\text{N}$ : C, 58.8; H, 4.9; N, 4.0;  $\text{CH}_3\text{O}$ , 17.9% (one methoxyl and one ethoxyl calculated as methoxyl); mol. wt., 347.

In view of the vigorous evolution of gas at the melting point, the acid was heated at 190–200 °C. for 15 minutes. The residue was dissolved in bicarbonate, filtered, and acidified, giving an immediate crystalline precipitate. This was recrystallized from ethyl acetate and obtained as colourless leaflets, m.p. 195–197 °C., not depressed by admixture with the acid by-product (m.p. 195–197 °C.) from the action of aqueous ethanolic potash on melicopidine (*vide supra*).

Found: C, 63.1; H, 5.4; N, 5.7;  $\text{CH}_3\text{O}$ , 12.5%.

Calculated for  $\text{C}_{13}\text{H}_{13}\text{O}_4\text{N}$ : C, 63.2; H, 5.3; N, 5.7;  $\text{CH}_3\text{O}$ , 12.6% (one methoxyl).

The filtrate from this acid was repeatedly extracted with chloroform, the chloroform evaporated and the residue crystallized from ethyl acetate. It separated as fine white needles, m.p. 181–182 °C.

Found: C, 63.6; H, 5.7; N, 4.8;  $\text{CH}_3\text{O}$ , 20.3%.

Calculated for  $\text{C}_{16}\text{H}_{17}\text{O}_5\text{N}$ : C, 63.4; H, 5.6; N, 4.6;  $\text{CH}_3\text{O}$ , 20.5% (two alkoxy groups calculated as methoxyl).

\* See footnote, p. 258.



*(h) Oxidation of the Dihydroxy Compounds*

(i) The dihydroxy compound obtained from melicopine by treatment with ethanolic potash followed by alcoholic hydrochloric acid is unstable in alkaline solution in contact with air. Oxidation was best effected by passing oxygen through the alkaline solution, when the colour changed from violet to pale reddish-brown. After saturating with carbon dioxide and extracting with chloroform, the solution was acidified and heated, giving a voluminous orange crystalline precipitate which proved to be the acid,  $C_{13}H_9O_4N$ , described above, m.p. and mixed m.p. 261–263 °C. (uncorr.). Yield 50–60%.

(ii) The dihydroxy compound from melicopidine and ethanolic potash followed by alcoholic acid was likewise oxidized in alkaline solution by oxygen. Acidification afforded a buff coloured precipitate which was purified by crystallization from dilute hydrochloric acid and finally from acetic acid. It was obtained as small colourless needles, m.p. 237–238 °C. (decomp.). Yield 65–70%.

Found: C, 56.6; H, 3.4; N, 5.1;  $CH_3O$ , nil;  $CH_3N$ , 11.0%; eq. wt. (potentiometric titration), 143.

Calculated for  $C_{13}H_9O_6N$ : C, 56.7; H, 3.3; N, 5.1;  $CH_3N$ , 10.5%; mol. wt. 275.

On boiling the filtrate from this acid, a small amount of the orange acid,  $C_{13}H_9O_4N$  (m.p. and mixed m.p.) separated. The filtrate from the orange acid was cooled and shaken with a small volume of chloroform. This extract deposited colourless crystals of a third acid, m.p. 254–256 °C. (uncorr.) after sintering at 245 °C.

The dimethoxydihydroxy and the isopropoxymethoxydihydroxy compounds from melicopidine were likewise oxidized in alkaline solution by oxygen to the same acid,  $C_{13}H_9O_6N$  (m.p. and mixed m.p.). In the former case a small amount of the third acid was also isolated, m.p. 252–254 °C. (uncorr.) after sintering at 245 °C.

## III. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Industrial Chemistry, C.S.I.R.O. The authors are indebted to R. J. L. Martin for carrying out potentiometric titrations.

## IV. REFERENCES

- (1) PRICE, J. R.—*Aust. J. Sci. Res. A* **2**: 249 (1949).
- (2) ROBINSON, G. M., and ROBINSON, R.—*J. Chem. Soc.* **111**: 929 (1917).

# ALKALOIDS OF THE AUSTRALIAN RUTACEAE: *MELICOPE* *FAREANA*

## III. OXIDATIVE DEMETHYLATION OF MELICOPICINE

By W. D. CROW\*

[Manuscript received March 9, 1949]

### Summary

Melicopine and *normelicopine* undergo oxidative demethylation with nitric acid giving rise to isomeric dimethoxyquinones. These are further demethylated by alkali to the same hydroxymethoxyquinone which is converted by hydrobromic acid to a dihydroxyquinone. All the quinones can be reductively methylated back to melicopine.

### I. INTRODUCTION

In the preceding paper of this series(1), the alkaloid melicopine,  $C_{18}H_{19}O_5N$ , was shown to be easily demethylated by alcoholic hydrochloric acid with the formation of *normelicopine*,  $C_{17}H_{17}O_5N$ . Both melicopine and *normelicopine* react vigorously with nitric acid, the same products being obtained from each. The main product is a bright red crystalline solid, m.p. 200.5-201.5 °C., having the empirical formula  $C_{16}H_{13}O_5N$  and containing two methoxyl groups. It is insoluble in saturated bisulphite solution. A second product, isomeric with the first, and likewise containing two methoxyl groups, is a dark red solid, m.p. 233-235 °C., soluble in saturated bisulphite solution. Both substances give 2,4-dinitrophenylhydrazones, and are readily reduced by sulphur dioxide to orange-red dihydro- compounds,  $C_{16}H_{15}O_5N$ , which give monoacetates with acetic anhydride and pyridine, and diacetates with acetic anhydride and anhydrous sodium acetate. The monoacetates of the dihydro- compounds may also be prepared by reductive acetylation of the parent substances. Further, the dihydro- compounds are reoxidized to the parent substances by nitric acid. The compound of m.p. 233-235 °C. (bisulphite-soluble) reacts rapidly with *o*-phenylenediamine to give a yellow phenazine,  $C_{22}H_{17}O_3N_3$ , whereas, under the same conditions its isomer is unchanged.

These reactions clearly show that the two compounds are isomeric quinones, the one of m.p. 233-235 °C. being an *o*-quinone (II) and its isomer a *p*-quinone (I), or possibly a bicyclic quinone. The formation of quinones by oxidative demethylation has been known for many years, and several instances are cited by Nightingale(2). Recently, Seshadri and his collaborators have made good use of the reaction in determining the structures of naturally occurring flavones and related compounds, for example, see Rao, Rao, and Seshadri(3).

\* Division of Industrial Chemistry, C.S.I.R.O.

The quinones do not give phenylhydrazones with phenylhydrazine, but, as happens frequently with benzoquinones, are reduced to quinols (III and IV). These are also found as by-products in the action of nitric acid on melicopicine or *normelicopicine*. The carbon skeleton of melicopicine is still present in the dimethoxyquinones, since both quinols are remethylated by dimethyl sulphate to *normelicopicine* and thence to melicopicine. The formation of the quinones from *normelicopicine*, with the loss of one methoxyl group, shows that the hydroxyl group in *normelicopicine* is converted to a quinonoid keto-group common to both quinones. This hydroxyl group is therefore phenolic.

The dimethoxyquinones may be demethylated step by step to a dihydroxyquinone. The first step in the demethylation is accomplished by heating with aqueous sodium carbonate, whereby the dimethoxyquinones both give purplish-red solutions, which yield the same red acid,  $C_{15}H_{11}O_5N$ , on acidification. This acid (V) is a hydroxymethoxyquinone. It contains one methoxyl group and gives a bright red *monoacetate* with acetic anhydride and pyridine. Reductive acetylation yields a *diacetyldihydro-* derivative and this can be further acetylated by acetic anhydride and anhydrous sodium acetate to a *triacetyldihydro-* derivative. The acid is reduced by sulphur dioxide to a *dihydro-* compound,  $C_{15}H_{13}O_5N$  (VI), which gives the same diacetyl derivative as is formed from the parent acid by reductive acetylation. The dihydro-compound is reconverted to the acid by aerial oxidation in alkaline solution. The action of sodium carbonate on the dimethoxyquinones is evidently an instance of the normal hydrolysis of hydroxyquinone ethers (e.g. 2-methoxy-1,4-naphthaquinone). This is confirmed by remethylation of the dihydro-compound to melicopicine. The hydroxymethoxyquinone condenses with *o*-phenylenediamine forming a *phenazine*,  $C_{21}H_{15}O_3N$ , which gives a *monoacetate* with acetic anhydride and pyridine. Acetylation under these conditions suggests that the acylable centre is not that present in *normelicopicine*. The same phenazine is obtained by the prolonged action of *o*-phenylenediamine on the dimethoxyquinone (I). Complete demethylation is attained when the hydroxymethoxyquinone is boiled with 46 per cent. hydrobromic acid. The *dihydroxyquinone* formed,  $C_{14}H_9O_5N$  (VII), is a dark red microcrystalline substance. It contains one methylimino group but no methoxyl and may be reduced by sulphur dioxide to a *dihydro-* compound, which is acetylated by acetic anhydride and pyridine to a triacetate. The *triacetate* is also obtained by reductive acetylation of the dihydroxyquinone. Both direct acetylation and condensation with *o*-phenylenediamine lead to mixtures which are not separated by crystallization.

Reductive methylation of the dihydroxyquinone gives melicopicine, which can therefore be demethylated in four stages without change in the fundamental structure of the molecule. These stages and the interrelations of the compounds described may be represented as shown in Figure 1. The partial structures shown provide the most plausible interpretation of the facts, though there is as yet no proof that the four methoxyl groups of melicopicine are located in the same ring. The conversion of melicopicine and *normelicopicine* to the dimethoxyquinones (I and II) can be effected by nitrous acid as well as by nitric. An

analogy to this reaction is provided by the conversion of 2,4,5-trimethoxybenzoic acid to the oxime of 4,5-dimethoxy-1,2-benzoquinone by nitrous acid(4).

The dihydroxyquinone (VII) is very easily oxidized. On shaking an alkaline solution in air, the green colour is destroyed and acidification and boiling yield the orange acid,  $C_{13}H_9O_4N(1)$ . When the dihydroxyquinone is

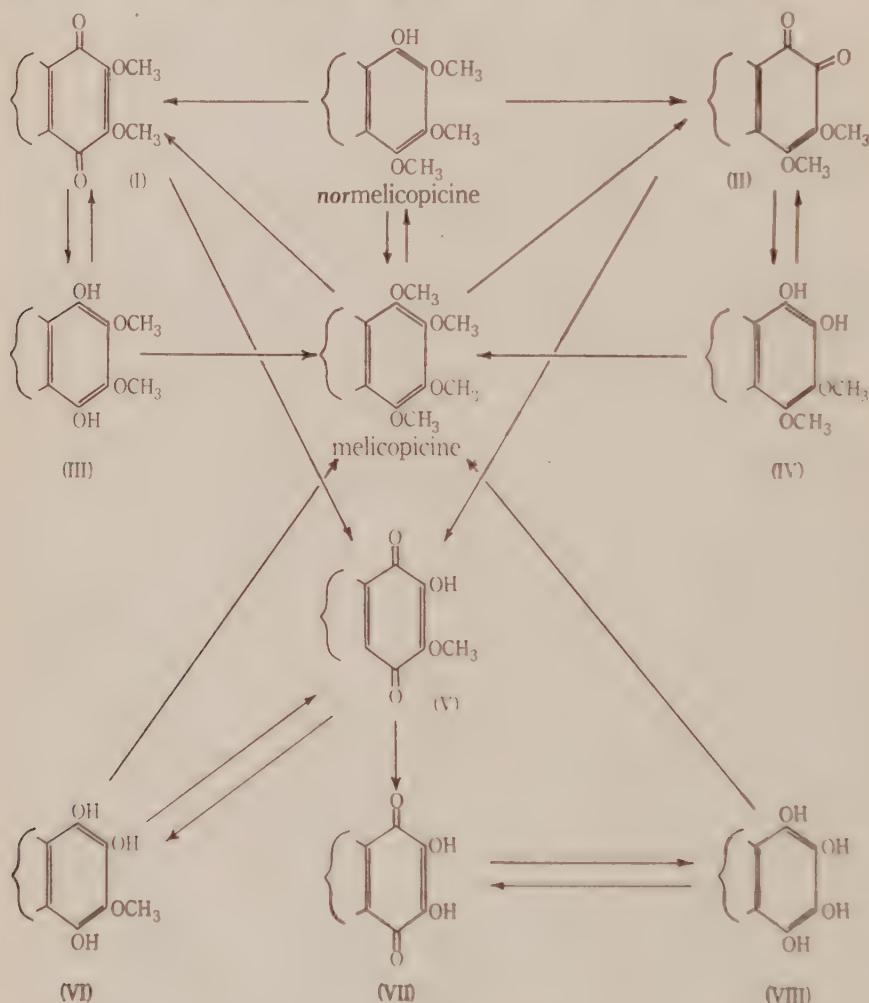
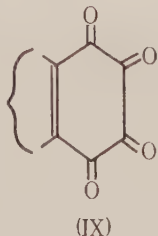


Fig. 1

boiled with water disproportionation takes place and the greater part is reduced to the quinol. The hydroxymethoxyquinone (V) which is stable in boiling alkaline (10 per cent. caustic soda) or acid solutions, behaves likewise. On heating with water it is slowly converted to the dihydro- compound, some of the orange acid,  $C_{13}H_9O_4N$ , being formed at the same time. The dihydroxyquinone is oxidized by cold nitric acid giving an unstable cream solid,  $C_{14}H_{11}O_7N$ . The reaction resembles that of nitric acid on *isonaphthazarin* and hexahydroxybenzene



which give tetraketotetralin(5) and hexaketocyclohexane(6) respectively. The pronounced tendency of such polyketo- compounds to form hydrates makes it probable, in the light of the partial structure suggested for the dihydroxyquinone, that the substance isolated is the dihydrate of a tetraketone (IX),  $C_{14}H_7O_5N \cdot 2H_2O$ .



This structure is in accord with the physical properties as well as with the mode of formation. It does not account satisfactorily for the failure of attempts to reduce the substance to the dihydroxyquinol. Moreover, the compound dissolves in caustic soda giving a green solution from which the acid,  $C_{13}H_9O_4N$ , may be obtained by aerial oxidation, followed by acidification and boiling. This reaction does not proceed in the absence of oxygen, though it appears to be analogous to the formation of croconic acid from hexaketocyclohexane which is formulated as a benzilic acid rearrangement.

## II. EXPERIMENTAL

All melting points are corrected unless otherwise stated. Microanalyses were carried out by R. B. Bradbury and H. L. Oates.

### (a) Preparation of the Dimethoxyquinones

(i) *Action of Nitric Acid on Melicopicine*.—Melicopicine (10 g.) was dissolved in 15% nitric acid (50 ml.) and 68% nitric acid (10 ml.) added to initiate the reaction. After the vigorous effervescence ceased (1–2 minutes) the mixture was poured onto excess sodium bicarbonate, diluted to 300 ml., and extracted repeatedly with chloroform. The combined chloroform extracts were concentrated to c. 50 ml. Ether (200 ml.) was added, precipitating a red solid (4–5 g.) which was separated into two constituents (I and II) by fractional crystallization from benzene. The yield of I was 40–50%, that of II was variable, not exceeding 8%, and occasionally negligible. The chloroform-ether filtrate was evaporated to dryness and the residue (1–1½ g.) fractionally crystallized from ethanol, giving III and, in greater amount, the more soluble IV.

The bicarbonate solution, after extraction of the above materials, was acidified and again extracted with chloroform. The aqueous acid solution was then concentrated to c. 200 ml. and allowed to stand, when a colourless acid separated. This melted at 237–238 °C. (decomp.), not depressed by admixture with the acid,  $C_{13}H_9O_6N$ , described in Part II(1). Yield 15–25%.

Found: C, 56.9; H, 3.3; N, 5.2%.

Calculated for  $C_{13}H_9O_6N$ : C, 56.7; H, 3.3; N, 5.1%.

The chloroform extracts were shaken first with aqueous sodium bicarbonate (5–10 ml.) and then with caustic soda. The former on acidification gave V. The caustic soda solution was yellow with a strong green fluorescence and acidification precipitated an orange solid, m.p. 260–262 °C. (uncorr.) mixed m.p. with the acid,  $C_{13}H_9O_4N$ (1), 261–263 °C. (uncorr.).

(ii) *Action of Nitric Acid on norMelicopicine*.—norMelicopicine was treated with cold 68% nitric acid, the reaction mixture diluted with water, poured onto excess sodium bicarbonate, and

extracted with chloroform. The chloroform extract yielded I and II, identified by m.p. and mixed m.p.

(iii) *Action of Nitrous Acid on norMelicopicine.*—*norMelicopicine* was triturated with 35% hydrochloric acid (20 ml./g.) and treated with a solution of nitrous acid (1 mole) at 0 °C. The reaction mixture was worked up as before and the same mixture of dimethoxyquinones isolated. The hydroxymethoxyquinone (V), identified by m.p. and mixed m.p., was formed in small amount, but no quinols could be detected.

(iv) *Action of Nitrous Acid on Melicopicine.*—*Melicopicine* was not attacked by one mole of nitrous acid, but the mixture of I and II resulted when it was treated in 35% hydrochloric acid with excess (8–10 moles) nitrous acid. The acid,  $C_{13}H_9O_6N$ , was obtained as a by-product.

#### (b) Properties of the Dimethoxyquinones

The compound (I) crystallized from methanol as bright red needles, m.p. 200.5–201.5 °C.

Found: C, 64.1; H, 4.4; N, 4.9;  $CH_3O$ , 20.7%.

The isomer (II) crystallized from methanol as dark red needles, m.p. 233–235 °C.

Found: C, 64.1; H, 4.3; N, 4.8;  $CH_3O$ , 20.4%.

Calculated for  $C_{16}H_{13}O_5N$ : C, 64.2; H, 4.3; N, 4.7;  $CH_3O$ , 20.7% (two methoxyls).

I and II were readily reduced by sulphur dioxide in methanol solution to the *dihydro*-compounds, III and IV respectively. III separated from ethanol as orange-red needles, m.p. 243–244 °C.

Found: C, 63.8; H, 5.1; N, 4.8%.

IV crystallized from aqueous ethanol as orange-red needles, m.p. 163.5–164.5 °C.

Found: C, 63.6; H, 5.1; N, 4.8%.

Calculated for  $C_{16}H_{15}O_5N$ : C, 63.8; H, 5.0; N, 4.7%.

The specimens of III and IV obtained as by-products in the action of nitric acid on *melicopicine* were identified by m.p.'s and mixed m.p.'s of the quinols themselves and of their monoacetates (*vide infra*). The two quinols dissolved in caustic soda, giving red solutions from which they were precipitated by carbon dioxide. Oxidation with cold 68% nitric acid reconverted them to the respective quinones (I and II; m.p. and mixed m.p.). Reductive acetylation of I and II with acetic anhydride and zinc dust gave *monoacetyldihydro*-compounds identical with monoacetates obtained by direct acetylation of III and IV respectively with acetic anhydride and pyridine. The *monoacetyldihydro*-derivative of I separated as orange-yellow needles from ethyl acetate-light petroleum, m.p. 147–148 °C.

Found: C, 63.0; H, 5.3; N, 4.1;  $CH_3CO$ , 13.1%.

The corresponding derivative of II was obtained as orange-yellow needles from ethyl acetate-light petroleum, m.p. 158–159 °C.

Found: C, 63.1; H, 5.2; N, 4.3;  $CH_3CO$ , 12.6%.

Calculated for  $C_{16}H_{14}O_4N(OOC.CH_3)$ : C, 63.0; H, 5.0; N, 4.1;  $CH_3CO$ , 12.5%.

Both these monoacetates were further acetylated to *diacetates* by refluxing with acetic anhydride and anhydrous sodium acetate. The *diacetyldihydro*-derivative of I crystallized from chloroform-light petroleum as pale yellow prisms, m.p. 176–177 °C.

Found: C, 62.5; H, 5.0; N, 3.7;  $CH_3CO$ , 22.3%.

That derived from II, pale yellow prisms from chloroform-light petroleum, melted at 204–204.5 °C.

Found: C, 62.4; H, 4.8; N, 3.7;  $CH_3CO$ , 22.2%.

Calculated for  $C_{16}H_{13}O_3N(OOC.CH_3)_2$ : C, 62.3; H, 4.9; N, 3.6;  $CH_3CO$ , 22.3%.

The *dihydro*-compounds were methylated by dimethyl sulphate and caustic soda containing a little sodium hydrosulphite to prevent oxidation. In each case the first product was *nor-melicopicine* (m.p. and mixed m.p.) which was further methylated to *melicopicine*, the second methyl group being introduced more slowly.

Both dimethoxyquinones reacted with Brady's reagent giving dark red *2,4-dinitrophenylhydrazones*; that from I, needles from chloroform-ether, melted at 255–256 °C. (decomp.).

Found: C, 54.9; H, 3.3%.

The *2,4-dinitrophenylhydrazone* from II, needles from chloroform-alcohol, melted at 260–261 °C. (uncorr.).

Found: C, 54.9; H, 3.6%.

Calculated for  $C_{22}H_{17}O_8N_5$ : C, 55.1; H, 3.5%.

Phenylhydrazones were not obtained when the dimethoxyquinones were treated with alcoholic phenylhydrazine hydrochloride. Instead, both quinones were reduced to the respective quinols (m.p. and mixed m.p.). The quinol obtained in this way from I was analysed.

Found: C, 64.1; H, 5.0; O, 26.3(7); N, 4.7%.

Calculated for  $C_{16}H_{18}O_5N$ : C, 63.8; H, 5.0; O, 26.6; N, 4.7%.

The dimethoxyquinone (II) reacted rapidly in alcoholic solution at room temperature with *o*-phenylenediamine (1.2 moles) with the formation of a *phenazine*, yellow needles from ethyl acetate-light petroleum, m.p. 244–245 °C.

Found: C, 71.2; H, 4.5; N, 11.4;  $CH_3O$ , 15.9%.

Calculated for  $C_{22}H_{17}O_3N_3$ : C, 71.2; H, 4.6; N, 11.3;  $CH_3O$ , 16.7% (two methoxyls).

Under the same conditions, the isomeric quinone (I) did not react and was recovered unchanged after standing with *o*-phenylenediamine for 24 hours (m.p. and mixed m.p.). The inference that II is an *o*-quinone was supported by the observation that it dissolved in saturated aqueous bisulphite while I did not.\*

#### (c) Preparation of the Hydroxymethoxyquinone (V)

The dimethoxyquinones (I and II) were insoluble in cold aqueous sodium carbonate but when heated on the water-bath both dissolved, giving purplish-red solutions. On acidification the same *acid* (m.p. and mixed m.p.) was obtained from both, as a crystalline precipitate. This substance (V) was identical (m.p. and mixed m.p.) with the specimen obtained as a by-product in the action of nitric acid on melicopicine. It crystallized from aqueous methanol as red needles, m.p. 247–248 °C. (decomp.).

Found: C, 63.2; H, 4.1; N, 5.0;  $CH_3O$ , 10.8%.

Calculated for  $C_{16}H_{11}O_5N$ : C, 63.2; H, 3.9; N, 4.9;  $CH_3O$ , 10.9% (one methoxyl).

Reduction of the hydroxymethoxyquinone in alcoholic solution by sulphur dioxide gave the *dihydro*- compound, orange-red needles from aqueous methanol, m.p. 276–277 °C. (uncorr.).

Found: C, 63.1; H, 4.7; N, 4.9%.

Calculated for  $C_{16}H_{13}O_5N$ : C, 62.7; H, 4.5; N, 4.9%.

The *dihydro*- compound dissolved in caustic soda giving an intense violet solution which underwent aerial oxidation and on acidification afforded the parent hydroxymethoxyquinone (m.p. and mixed m.p.). Acetylation of the quinone with acetic anhydride and pyridine gave the *monoacetate*, crimson needles from ethyl acetate, m.p. 236–237 °C. (decomp.).

Found: C, 62.2; H, 4.1; N, 4.4;  $CH_3CO$ , 13.5%.

Calculated for  $C_{16}H_{10}O_4N(OOC.CH_3)$ : C, 62.4; H, 4.0; N, 4.3;  $CH_3CO$ , 13.1%.

Reductive acetylation of this acetyl derivative gave the same *diacetyldihydro*- compound as was obtained by reductive acetylation of the hydroxymethoxyquinone and by direct acetylation of the quinol. The product crystallized from ethyl acetate-light petroleum as yellow needles, m.p. 214–215 °C.

Found: C, 61.7; H, 4.6; N, 4.0;  $CH_3CO$ , 23.5%.

Calculated for  $C_{16}H_{11}O_3N(OOC.CH_3)_2$ : C, 61.5; H, 4.6; N, 3.8;  $CH_3CO$ , 23.2%.

\* This property could not be utilized in the separation of the two quinones because of difficulties in recovering II from the bisulphite.



Further acetylation by acetic anhydride and anhydrous sodium acetate yielded the *triacetyl-dihydro-* compound, pale yellow prisms from ethyl acetate-light petroleum, m.p. 211–213 °C.

Found: C, 61.0; H, 4.6; N, 3.6;  $\text{CH}_3\text{O}$ , 7.3;  $\text{CH}_3\text{CO}$ , 31.0%.

Calculated for  $\text{C}_{15}\text{H}_{10}\text{O}_2\text{N}(\text{OOC}\cdot\text{CH}_3)_3$ : C, 61.0; H, 4.6; N, 3.4;  $\text{CH}_3\text{O}$ , 7.5;  $\text{CH}_3\text{CO}$ , 31.2%.

Methylation of the dihydro- compound by dimethyl sulphate and caustic soda containing some sodium hydrosulphite gave melicopicine, identified by m.p. and mixed m.p. and confirmed by conversion to *normelicopicine* with alcoholic hydrochloric acid. The hydroxymethoxyquinone reacted with *o*-phenylenediamine in alcoholic solution. The mixture was allowed to stand at room temperature for several days, during which the *phenazine* slowly crystallized. After recrystallization from ethyl acetate-chloroform it was obtained as dark violet needles, m.p. 258–259 °C.

Found: C, 70.5; H, 4.3%.

Calculated for  $\text{C}_{21}\text{H}_{15}\text{O}_3\text{N}_3$ : C, 70.6; H, 4.2%.

The same phenazine derivative (m.p. and mixed m.p.) resulted from the prolonged action of alcoholic *o*-phenylenediamine on the dimethoxyquinone (I). Acetylation of the phenazine by acetic anhydride and pyridine gave a *monoacetate*, deep yellow prisms from chloroform-light petroleum, m.p. 250–251 °C.

Found: C, 69.7; H, 4.6; N, 10.5;  $\text{CH}_3\text{CO}$ , 11.2%.

Calculated for  $\text{C}_{21}\text{H}_{14}\text{O}_2\text{N}_3(\text{OOC}\cdot\text{CH}_3)$ : C, 69.2; H, 4.3; N, 10.5;  $\text{CH}_3\text{CO}$ , 10.8%.

When the hydroxymethoxyquinone was boiled with water for an hour, the mixture darkened considerably and the bulk of the material became insoluble in sodium bicarbonate. The filtered solid was dried and identified as the quinol (VI) by conversion to the diacetate. The bicarbonate filtrate was treated with sodium hydrosulphite, the precipitate filtered and discarded, and the filtrate acidified. The orange precipitate after crystallization from acetic acid was identified (m.p. and mixed m.p.) as the acid,  $\text{C}_{13}\text{H}_9\text{O}_4\text{N}$ , described in Part II(1).

The hydroxymethoxyquinone was recovered unchanged (identified by m.p. and mixed m.p. of the diacetyldihydro- compound) after refluxing with 10% caustic soda for 6 hours.

#### (d) Preparation of the Dihydroxyquinone (VII)

The hydroxymethoxyquinone (3.0 g.) was refluxed with 46% hydrobromic acid (150 ml.). The hydrobromide of the hydroxymethoxyquinone, which precipitated immediately, was replaced during the refluxing by the insoluble hydrobromide of the demethylation product. After 3 hours the mixture was cooled, the hydrobromide filtered and washed with cold water until completely hydrolysed as shown by absence of bromide ion in the filtrate. The dark red microcrystalline residue had no definite melting point, but decomposed above 250 °C. It was analysed without further purification.

Found: C, 62.3; H, 3.5; N, 5.1%.

Calculated for  $\text{C}_{14}\text{H}_9\text{O}_2\text{N}$ : C, 62.0; H, 3.3; N, 5.2%.

The *dihydroxyquinone* was analysed without further purification because it could not be crystallized satisfactorily: for example, it undergoes disproportionation when heated with water (*vide infra*).

Acetylation of the dihydroxyquinone by acetic anhydride and pyridine at room temperature yielded a red solid which remained sticky after repeated attempts at crystallization, and was evidently a mixture. The quinone also reacted readily with *o*-phenylenediamine in alcohol at room temperature giving a black solid, which, like the acetylation product, could not be purified. The dihydroxyquinol was prepared by reduction of the quinone and was purified for analysis through its triacetate. The quinone hydrobromide was suspended in sodium bicarbonate solution and reduced by sodium hydrosulphite. The product was filtered and washed, first with sodium bicarbonate solution, then with dilute potassium bisulphite, and finally with water. The dihydroxyquinol was acetylated with acetic anhydride and pyridine, giving the *triacetate*, yellow prisms from ethyl acetate-light petroleum, m.p. 193–194 °C.



Found: C, 60.4; H, 4.4; N, 3.6;  $\text{CH}_3\text{CO}$ , 32.3%.

Calculated for  $\text{C}_{14}\text{H}_8\text{O}_2\text{N}(\text{OOC}\cdot\text{CH}_3)_3$ : C, 60.2; H, 4.3; N, 3.5;  $\text{CH}_3\text{CO}$ , 32.3%.

The same triacetate (m.p. and mixed m.p.) was obtained directly by reductive acetylation of the dihydroxyquinone. The dihydroxyquinol was recovered from the purified triacetate by acid hydrolysis. The triacetate (1.0 g.) was dissolved in methanol (50 ml.) containing concentrated hydrochloric acid (5 ml.). After boiling 5 minutes, hot water (20 ml.) was added, and heating continued until the bulk of the methanol had evaporated. The dihydroxyquinol separated on cooling as orange-red needles which melted with decomposition at 304–306 °C. (uncorr.).

Found: C, 61.7; H, 4.1; N, 5.3;  $\text{CH}_3\text{O}$ , nil;  $\text{CH}_3\text{N}$ , 10.2%.

Calculated for  $\text{C}_{14}\text{H}_{11}\text{O}_6\text{N}$ : C, 61.5; H, 4.0; N, 5.1;  $\text{CH}_3\text{N}$ , 10.6%.

Methylation of the quinol, or of its triacetate, with dimethyl sulphate and caustic soda containing sodium hydrosulphite gave melicopicine, identified by m.p. and mixed m.p.

The quinol dissolved in caustic soda giving a red solution which on shaking underwent aerial oxidation, the colour changing to olive-green and finally to pale reddish-brown. The quinone itself gave a sparingly soluble salt with sodium bicarbonate, the solution being green. With caustic soda it dissolved to an olive-green solution, which, as with the quinol, underwent aerial oxidation changing to pale reddish-brown. Acidification and heating afforded the orange acid,  $\text{C}_{13}\text{H}_9\text{O}_4\text{N}$ , m.p. and mixed m.p. 261–262 °C. (uncorr.). The quinone underwent disproportionation on boiling with water. The insoluble product was filtered off and identified as the dihydroxyquinol (VIII) by the red colour of a caustic soda solution and by direct acetylation to the triacetate (m.p. and mixed m.p. 192–194 °C.).

*The Action of Nitric Acid on the Dihydroxyquinone.*—The dihydroxyquinone was dissolved in the minimum amount of cold 68% nitric acid. When effervescence ceased, the pale yellow solution was diluted with water (10 vol.), and on standing, pale yellow needles separated. These were filtered as soon as possible, washed with water and acetone, and dried *in vacuo* at room temperature. The substance melted at 263–266 °C. (decomp., uncorr.) when placed in a bath preheated to 260 °C. Otherwise it decomposed gradually above 200 °C.

Found: C, 55.7; H, 3.8; N, 4.7%.

Calculated for  $\text{C}_{14}\text{H}_{11}\text{O}_7\text{N}$ : C, 55.1; H, 3.6; N, 4.6%.

The substance could not be recrystallized as it separated from water, the only suitable solvent, as a gel. The substance dissolved in caustic soda to give an emerald-green solution which underwent aerial oxidation on shaking. The resulting colourless solution when acidified and heated, precipitated orange needles of the acid,  $\text{C}_{13}\text{H}_9\text{O}_4\text{N}$ , m.p. and mixed m.p. 262–263 °C. Reduction with sulphur dioxide, tin and hydrochloric acid, zinc dust and caustic soda, or alkaline hydrosulphite did not give the dihydroxyquinone or, in fact, any satisfactory product. Similarly, reductive acetylation with zinc dust and acetic anhydride failed to yield an acetate of the quinol (VIII)—the product was a red gum.

### III. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Industrial Chemistry, C.S.I.R.O. The author is indebted to Dr. J. R. Price for helpful discussion and advice during the investigation.

### IV. REFERENCES

- (1) CROW, W. D., and PRICE, J. R.—*Aust. J. Sci. Res. A* **2**: 255 (1949).
- (2) NIGHTINGALE, D. V.—*Chem. Rev.* **40**: 117 (1947).
- (3) RAO, G. S. K., RAO, K. V., and SESHADRI, T. R.—*Proc. Ind. Acad. Sci.* **27A**: 245 (1948).
- (4) FABINYI, R., and SZEKI, T.—*Ber. deutsch. chem. Ges.* **39**: 3679 (1906).
- (5) ZINCKE, T., and OSSENBECK, A.—*Liebigs Ann.* **307**: 4 (1899).
- (6) NIETZKI, R., and BENCKISER, T.—*Ber. deutsch. chem. Ges.* **18**: 499 (1885).
- (7) UNTERZAUCHER, J.—*Ibid.* **73**: 391 (1940).

# ALKALOIDS OF THE AUSTRALIAN RUTACEAE: *MELICOPE FAREANA*

## IV. SOME REACTIONS OF 1-METHYL-4-QUINOLONE-3-CARBOXYLIC ACID, A DEGRADATION PRODUCT OF THE ALKALOIDS

By J. R. PRICE\*

[Manuscript received March 9, 1949]

### Summary

Melicopine, melicopidine, and melicopicine are each oxidized by nitric acid to the same acid,  $C_{11}H_9O_3N$ . Decarboxylation gives 1-methyl-4-quinolone (I) and the acid is shown to be 1-methyl-4-quinolone-3-carboxylic acid (II). The 3- and 6- mononitro-, 3,6-dinitro-, and 2-hydroxy-3,6-dinitro- derivatives of 1-methyl-4-quinolone have been prepared and the orientation of the substituents established by oxidation of the hydroxydinitro- compound to 5-nitro-*N*-methylantranilic acid. 1-Methyl-4-quinolone-3-carboxylic acid is readily converted to a tetrahydro- derivative. The carboxyl groups in II and in the tetrahydro- acid are reduced by zinc and hydrochloric acid to methyl groups.

### I. INTRODUCTION

Crow(1) has shown that nitric acid converts melicopicine to a mixture of two isomeric dimethoxyquinones. Under more vigorous conditions, however, further oxidation gives rise to a colourless weak acid,  $C_{11}H_9O_3N$ , which may be obtained likewise from melicopine and melicopidine. This acid contains no methoxyl group, cannot be acetylated, and does not react with nitrous acid. It is stable to caustic soda at 200° C. and is not attacked by boiling hydriodic acid and red phosphorus. The presence of a carboxyl group is demonstrated by the formation of an ethyl ester and by decarboxylation to a substance,  $C_{10}H_9ON$ , a weak base containing one methylimino group. This base, like the parent acid, cannot be acetylated and does not react with 2,4-dinitrophenylhydrazine. Comparison with a synthetic specimen proves it to be 1-methyl-4-quinolone (I), which has been found in nature as the alkaloid echinopsine(2). The properties of the base and the melting points of the picrate, platinumchloride, and hydrochloride are in general agreement with those recorded by Spath and Kolbe(3), though, as shown in Table I, the observed figures are somewhat higher. Short treatment of the base with 68 per cent. nitric acid gives a mononitro- derivative identical with that prepared from synthetic 1-methyl-4-quinolone. Further confirmation of the structure is provided by high pressure hydrogenation which gives a base,  $C_{10}H_{19}N$ , the picrate of which melts at the temperature recorded by Ehrenstein and Bunge(4) for the picrate of *cis*-1-methyl-decahydroquinoline.

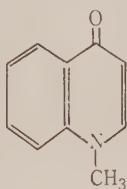
\* Division of Industrial Chemistry, C.S.I.R.O.

Adding bromine water to an aqueous solution of the base gives a dibromo-1-methyl-4-quinolone, whereas with alkaline hypobromite a monobromo-compound is obtained. The same monobromo-compound results, in low yield, when the acid,  $C_{11}H_9O_3N$ , is treated with hypobromite; compare the conversion of 4-hydroxyquinoline-3-carboxylamide to 3-bromo-4-hydroxyquinoline by hypobromite(5). By analogy with the 1-methyl-2- and 4-pyridones and 1-methyl-

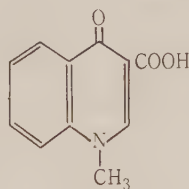
TABLE 1  
MELTING POINTS OF SALTS OF 1-METHYL-4-QUINOLONE

Salt	Observed Melting Point (°C., corr.)	Melting Point (°C., reported by 3)
Picrate ..	231-232	223-224
Platinichloride ..	219.5-220.5	210-212
Hydrochloride ..	196-198	185-186

2-quinolone, the mononitro- and monobromo- derivatives of 1-methyl-4-quinolone should be the 3-, 6-, or 8- compounds, and consequently the carboxyl group of the acid,  $C_{11}H_9O_3N$ , should be in the 3, 6, or 8 position. In fact, it is in the 3 position as is shown by comparison of the acid with a synthetic specimen of 1-methyl-4-quinolone-3-carboxylic acid (II).<sup>\*</sup> Additional evidence for structure II is furnished by the identity of the tetrahydro- acid,  $C_{11}H_{13}O_3N$ , obtained from both the synthetic and alkaloid-derived specimens of the acid by reduction with Raney alloy and caustic soda(6). Thus the three alkaloids appear to be derivatives of I with one point of attachment of the remainder of the molecule at position 3 of the quinolone nucleus.



(I)

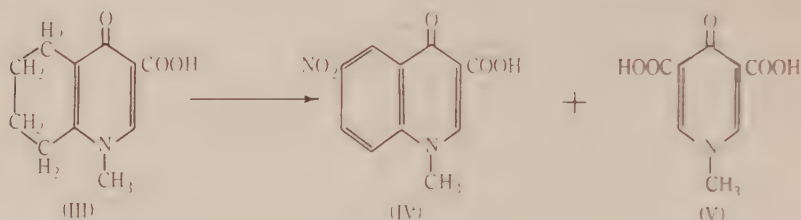


(II)

The tetrahydro- acid resulting from reduction of II is evidently 1-methyl-5,6,7,8-tetrahydro-4-quinolone-3-carboxylic acid (III). When refluxed with 68 per cent. nitric acid for several hours it is converted to a mixture of 6-nitro-1-methyl-4-quinolone-3-carboxylic acid (IV),<sup>\*</sup> and an acid,  $C_8H_7O_5N$ , still containing a methylimino group. This is presumably 1-methyl-4-pyridone-3,5-dicarboxylic acid (V), which is not recorded in the literature. Thus the nitric acid brings about dehydrogenation and nitration as the principal reactions, and to a lesser extent, oxidative ring fission. The tetrahydro- acid (III) is

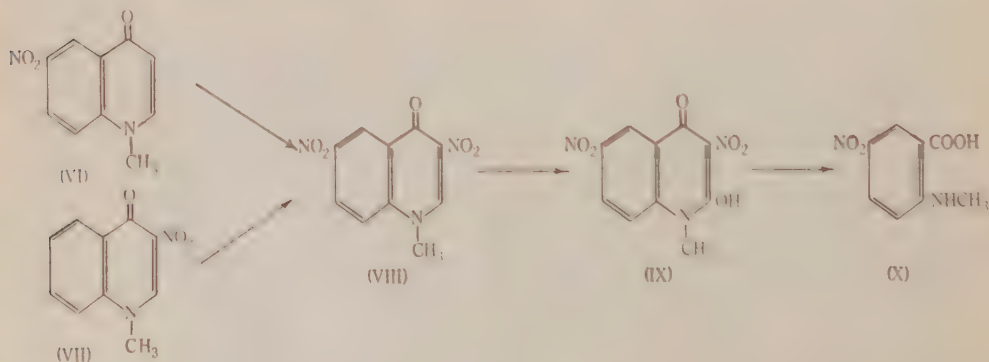
<sup>\*</sup> A description of the synthesis of these acids by L. J. Drummond and F. N. Lahey will be published shortly.

decarboxylated to a base,  $C_{10}H_{13}ON$  (1-methyl-5,6,7,8-tetrahydro-4-quinolone). This could not be prepared from 1-methyl-4-quinolone which was recovered unchanged after treatment with Raney alloy and caustic soda.



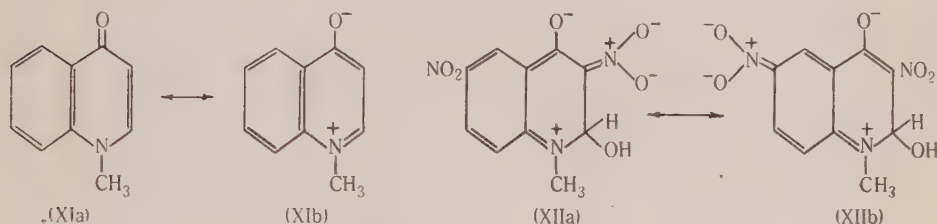
6-Nitro-1-methyl-4-quinolone-3-carboxylic acid (IV) is also formed by the prolonged action of nitric acid on II. It is decarboxylated to 6-nitro-1-methyl-4-quinolone (VI). This is not identical with the product of direct nitration of 1-methyl-4-quinolone, which should therefore be the 3- or 8-nitro-derivative. Both mononitro- compounds on further treatment with nitric acid give the same dinitro-1-methyl-4-quinolone which should therefore be either the 3,6- or the 6,8-dinitro- compound. Some of this dinitro- compound is formed by the action of nitric acid on II, together with the 3- (or 8)-mononitro- compound, picric acid,  $C_8H_7O_5N$  (V), and 6-nitro-1-methyl-4-quinolone-3-carboxylic acid (IV).

The dinitro- compound dissolves in caustic soda to an orange-yellow solution from which it is reprecipitated by carbon dioxide. In alkaline solution it is oxidized by permanganate to an acidic substance,  $C_{10}H_7O_2N(NO_2)_2$ , presumably a hydroxydinitro- compound. Further oxidation with nitric acid then yields 5-nitro-*N*-methylantranilic acid (X), proving that the second nitro- group is in the pyridone ring and that the hydroxyl group has also entered the pyridone ring. The mononitro- compound formed by direct nitration of I is therefore 3-nitro-1-methyl-4-quinolone (VII), the dinitro- compound is 3,6-dinitro-1-methyl-4-quinolone (VIII) and the hydroxy compound 2-hydroxy-3,6-dinitro-1-methyl-4-quinolone (IX).



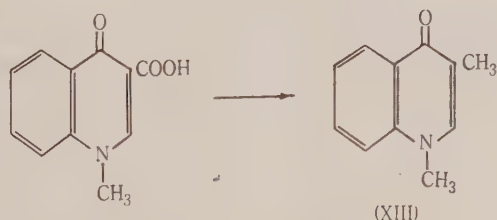


The solubility of the dinitro- compound in alkali is reminiscent of the behaviour of 1,3,5-trinitrobenzene in which the ion is stabilized by resonance between three quinonoid *aci*-nitro structures(7). 1-Methyl-4-quinolone itself is a weak base due to resonance between structures XIa and XIb. The ionic *aci*-nitro quinonoid structures (XIIa and XIIb) may be written for its dinitro- derivative, thus accounting for the solubility in alkali and the oxidation in alkaline solution to IX. Structures analogous to XIIa and XIIb cannot be written if the nitro-groups in the dinitro- compound are in the unlikely 2,6 positions.



In accordance with the orientation of the nitro- compounds, the monobromo-compound obtained from I should be 3-bromo-1-methyl-4-quinolone, in keeping with its formation by simultaneous bromination and decarboxylation of II. The dibromo- compound is presumably 3,6-dibromo-1-methyl-4-quinolone. 3-Amino-1-methyl-4-quinolone has been prepared by reduction of the nitro-compound.

1-Methyl-4-quinolone-3-carboxylic acid (II) is reduced by zinc dust and hydrochloric acid to a base,  $C_{11}H_{11}ON$ . The tetrahydro- acid (III) under the same conditions gives a base,  $C_{11}H_{15}ON$  (as shown by analysis of its picrate). The empirical formulae suggest reduction of the carboxyl group to methyl, and this is confirmed by comparison of the base,  $C_{11}H_{11}ON$ , with a synthetic specimen of 1,3-dimethyl-4-quinolone (XIII)\*. This remarkable reduction of a carboxyl group to methyl has only been observed in a few cases, e.g. the reduction of picolinic acid to  $\alpha$ -picoline(8) and of terephthalic acid to *p*-toluic acid(9).



## II. EXPERIMENTAL

All melting points are corrected unless otherwise stated. Microanalyses were carried out by R. B. Bradbury and H. L. Oates.

\* This synthesis, carried out by L. J. Drummond and F. N. Lahey, will be published shortly (cf. footnote p. 273).

(a) *Oxidation of the Alkaloids by Nitric Acid*

Melicopidine (2 g.) was treated cautiously with 68% nitric acid (10 ml.). After the violent initial reaction subsided, the greater part of the nitric acid was boiled off as quickly as possible. The brown viscous residue was dissolved in boiling acetic acid, the solution filtered, and allowed to crystallize. A further quantity of the oxidation product was obtained after concentration of the mother liquors. Recrystallization from acetic acid and finally from acetic anhydride afforded colourless needles which softened at c. 280 °C. and melted at 295–296 °C. (uncorr.) with decomposition.

Found: C, 65.2; H, 4.6; N, 7.0%;  $\text{CH}_3\text{O}$ , nil.

Calculated for  $\text{C}_{11}\text{H}_9\text{O}_3\text{N}$ : C, 65.0; H, 4.4; N, 6.9%.

The same acid (m.p. and mixed m.p.) was obtained from melicopine and melicopine. Yields under the above conditions were 60–70%, but considerably less was obtained from larger scale preparations, apparently due to the longer time of contact with boiling nitric acid. The substance is a weak acid, almost completely precipitated by carbon dioxide from alkaline solution. The *ethyl ester*, prepared by treatment of the acid with thionyl chloride followed by ethanol, in the usual manner, separated from benzene as colourless needles, m.p. 123–125 °C.

Found: C, 67.6; H, 5.5; N, 5.9%.

Calculated for  $\text{C}_{13}\text{H}_{13}\text{O}_3\text{N}$ : C, 67.5; H, 5.6; N, 6.1%.

The acid was recovered unchanged (m.p. and mixed m.p.) after treatment with the following reagents: (a) boiling acetic anhydride and anhydrous sodium acetate (this was carried out before the introduction of acetic anhydride as a solvent for recrystallizing the acid), (b) nitrous acid, (c) zinc dust and caustic soda, (d) anhydrous aluminium bromide, (e) boiling hydriodic acid and red phosphorus, and (f) 20% caustic soda at 200 °C.

It was only oxidized with difficulty by, for example, hot acid or alkaline permanganate or chromic anhydride in boiling acetic acid and in such cases no oxidation product could be isolated. The acid was found to be identical with a synthetic specimen of 1-methyl-4-quinolone-3-carboxylic acid (II), which melted at 296–297 °C. (uncorr.) after softening at c. 280 °C. A mixture of the two melted at 295–296 °C. (uncorr.) after softening at c. 280 °C. The acid was also identical (m.p. and mixed m.p.) with by-products from the action of ethanolic potash on melicopine or normelicopine, described in Part II(10).

(b) *Decarboxylation of the Acid,  $\text{C}_{11}\text{H}_9\text{O}_3\text{N}$* 

Decarboxylation could be effected in a variety of ways. Good yields of base were obtained in small scale experiments by heating the acid with zinc dust or the sodium salt with soda-lime. However, the most satisfactory procedure, particularly for larger quantities, consisted of heating the acid for 5 minutes, finally at the boiling point, with dibutyl phthalate containing a little copper-bronze. The solution was cooled, diluted with about 5 vol. chloroform, filtered, and extracted three times with 5% hydrochloric acid. The acid extracts were basified, shaken with chloroform, and the chloroform evaporated. The yield of crude base, somewhat discoloured but pure enough for most purposes, was c. 80%. Purification was best effected through the picrate or hydrochloride. The base, recovered from the recrystallized salt, was crystallized either from benzene or from a small volume of water. It separated from water as a hydrate, fine white needles which lost water on drying. The anhydrous base melted at 152–153 °C.

Found: C, 76.0; H, 5.7; N, 8.8;  $\text{CH}_3\text{O}$ , nil;  $\text{CH}_3\text{N}$ , 18.9%; mol. wt., 157 (modified Signer method).

Calculated for  $\text{C}_{10}\text{H}_9\text{ON}$ : C, 75.5; H, 5.7; N, 8.8;  $\text{CH}_3\text{N}$ , 18.2%; mol. wt. 159.

The base did not react with acetic anhydride and anhydrous sodium acetate, alcoholic 2,4-dinitrophenylhydrazine hydrochloride, or nitrous acid. Like the parent acid it was only attacked by strong oxidizing agents such as hot acid or alkaline permanganate, when no product could be isolated. It was recovered unchanged after treatment with permanganate in acetone or chromic anhydride in acetic acid at room temperature. It was identified as 1-methyl-4-quinolone (I) by comparison with a synthetic specimen. The melting point of the synthetic base,

alone or mixed with that prepared from the alkaloids, was 152–153 °C. Both specimens gave an identical blood-red colour with ferric chloride.\*

The base yields a *picrate*, sparingly soluble, small yellow prisms from methanol, m.p. 231–232 °C. alone or in admixture with the picrate of synthetic 1-methyl-4-quinolone.

Found: C, 49.8; H, 3.4; N, 14.5%.

Calculated for  $C_{10}H_9ON.C_6H_3O_7N_3$ : C, 49.5; H, 3.1; N, 14.4%

The *platinichloride* was obtained as dull golden-yellow plates from alcohol, m.p. 219.5–220.5 °C. (decomp.) alone or mixed with the platinichloride of the synthetic base. The hydrochloride was not satisfactory for comparative purposes. It is hygroscopic and loses hydrogen chloride on heating. However, a specimen dried under reduced pressure at 80 °C. over phosphorus pentoxide, melted at 195–197 °C. after previous softening. The hydrochloride of synthetic 1-methyl-4-quinolone, prepared and dried under the same conditions, had m.p. 196–198 °C. after previous softening. A mixture of the two melted at 195–197 °C.

The base obtained from the alkaloids was hydrogenated in alcohol over Raney nickel at 200 °C. and 900 lb./sq. in. The filtered solution was acidified, evaporated, the residue dissolved in water, and the solution treated with aqueous picric acid. The precipitated picrate, after repeated crystallization first from water, then from methanol, was obtained as yellow needles, m.p. 197–198 °C. (decomp.).

Found: C, 50.2; H, 5.8; N, 14.7%.

Calculated for  $C_{10}H_{19}N.C_6H_3O_7N_3$ : C, 50.3; H, 5.8; N, 14.7%.

Ehrenstein and Bunge(4) record 199–200 °C. as the m.p. of the picrate of *cis*-1-methyl-decahydroquinoline.

#### (c) Bromination of 1-Methyl-4-Quinolone

(i) Addition of excess bromine water to an aqueous solution of the base gave an immediate bright yellow voluminous precipitate. This was dissolved in chloroform, the solution washed with caustic soda (washings discarded), and evaporated. Repeated crystallization of the residue from methanol gave long colourless needles of the *dibromo*- compound, m.p. 249–251 °C. after sintering at 246 °C.

Found: C, 38.0; H, 2.4; N, 4.4; Br, 50.4%.

Calculated for  $C_{10}H_7ONBr_2$ : C, 37.9; H, 2.2; N, 4.4; Br, 50.5%.

(ii) The base (1.2 g.) dissolved in 5% caustic soda (20 ml.) was treated with sodium hypobromite (2.0 ml. bromine in 40 ml. 5% caustic soda), and allowed to stand overnight. The solid (1.04 g.) which separated was recrystallized from methanol and obtained as colourless needles, m.p. 233–235 °C.

Found: C, 50.2; H, 3.4; N, 6.0; Br, 33.6%.

Calculated for  $C_{10}H_8ONBr$ : C, 50.4; H, 3.4; N, 5.9; Br, 33.6%.

The same *monobromo*- compound resulted, in about 25% yield, from treatment of an alkaline solution of 1-methyl-4-quinolone-3-carboxylic acid (II) with sodium hypobromite. It melted at 233–235 °C. alone or mixed with the material prepared from the base.

#### (d) Nitro-derivatives of 1-Methyl-4-Quinolone

(i) *3-Nitro-1-Methyl-4-Quinolone* (VII).—The base prepared from the alkaloids was nitrated by boiling vigorously with 68% nitric acid for about 5 minutes. The solution was poured into water, basified with caustic soda, and extracted with chloroform. The residue after evaporation of the chloroform was crystallized from methanol, from which it separated as colourless needles, m.p. 227–229 °C. Yield variable, 50–80%.

Found: C, 58.7; H, 4.0; N, 13.8%.

Calculated for  $C_{10}H_8O_3N_2$ : C, 58.8; H, 3.9; N, 13.7%.

\* The colour is not particularly intense, though it is described as such by Greshoff(2).



3-Nitro-1-methyl-4-quinolone prepared in the same manner from synthetic 1-methyl-4-quinolone melted at 227–229 °C. alone or mixed with the material prepared from the alkaloids.

(ii) *3,6-Dinitro-1-Methyl-4-Quinolone (VIII)*.—The above mononitro- compound was refluxed with 68% nitric acid for 15 hours, the bulk of the nitric acid evaporated, and the residue basified with sodium bicarbonate. The filtered solid crystallized from acetic acid in short pale cream needles, m.p. 271–272 °C. (uncorr.). Yield, 70%.

Found: C, 48.2; H, 3.0; N, 17.0%.

Calculated for  $C_{10}H_7O_5N_2$ : C, 48.2; H, 2.8; N, 16.9%.

The substance is not soluble in aqueous sodium carbonate but dissolves in caustic soda to an orange-yellow solution from which it is precipitated by sodium bicarbonate. It is evidently a pseudo-acid as in some instances, acidification of a caustic soda solution with hydrochloric acid gave a red precipitate which reverted to cream on warming. Caustic soda solutions of the dinitro- compound decomposed on heating.

(iii) *2-Hydroxy-3,6-Dinitro-1-Methyl-4-Quinolone (IX)*.—The dinitro- compound, dissolved in 20% caustic soda (gentle warming) was oxidized by permanganate, and the manganese dioxide removed by acidification and addition of sodium sulphite. The cream precipitate which contained inorganic matter was triturated with concentrated hydrochloric acid, filtered, and crystallized from ethanol. It separated as colourless needles, m.p. 177–177.5 °C.

Found: C, 45.4; H, 2.7; N, 15.8%.

Calculated for  $C_{10}H_7O_6N_2$ : C, 45.3; H, 2.6; N, 15.8%.

The substance is almost insoluble in concentrated hydrochloric acid. It gives a sodium salt which is only sparingly soluble in water.

Further oxidation of the hydroxydinitro- compound was effected by boiling for 3–5 minutes with 68% nitric acid. The mixture was diluted with water, basified with bicarbonate, and extracted with chloroform. The chloroform extracts contained only a small amount of material which was not examined further. The deep yellow bicarbonate solution was acidified, giving a pale yellow precipitate, which crystallized from water as a dull yellow microcrystalline powder, m.p. 258–260 °C. (uncorr.).

Found: C, 48.7; H, 4.4; N, 14.0%.

Calculated for  $C_8H_8O_4N_2$ : C, 49.0; H, 4.1; N, 14.3%.

An authentic sample of 5-nitro-*N*-methylantranilic acid (X) melted at 259–261 °C. (uncorr.), the mixed m.p. was 259–260 °C. (uncorr.).

(e) *1-Methyl-5,6,7,8-Tetrahydro-4-Quinolone-3-Carboxylic Acid (III)*

(i) *Preparation of the Tetrahydro- Acid*. The acid (11: 2 g.) dissolved in 10% caustic soda (60 ml.), was heated on the water-bath, and reduced by the gradual addition of finely divided Raney alloy (6 g.) over a period of three hours. The cooled and filtered solution was acidified and extracted continuously with chloroform for 20 hours. The residue after removal of the chloroform crystallized from methanol as colourless needles, m.p. 275–277 °C. (uncorr.). Yield, 70–80%.

Found: C, 63.6; H, 6.2; N, 6.8%.

Calculated for  $C_{11}H_{13}O_3N$ : C, 63.8; H, 6.3; N, 6.8%.

The mixed melting point with a specimen prepared in the same manner from synthetic 1-methyl-4-quinolone-3-carboxylic acid was 275–277 °C. (uncorr.) (see footnote p. 277). 1-Methyl-4-quinolone was recovered unchanged (identified by m.p. and mixed m.p. of the picrate) after refluxing a solution of 1 g. in 10% caustic soda (35 ml.) and ethanol (20 ml.) for 5 hours with the gradual addition of Raney alloy (3.5 g.).

The tetrahydro- acid (III) was decarboxylated by heating for 3–5 minutes in dibutyl phthalate containing a little copper-bronze. The colourless crystalline base, 1-methyl-5,6,7,8-tetrahydro-4-quinolone, was not analysed as such, but was converted to the *picrate*, yellow prisms from methanol, m.p. 245–247 °C. (uncorr.) with decomposition.



Found: C, 49.3; H, 4.2; N, 14.3%.

Calculated for  $C_{10}H_{13}ON.C_6H_5O_7N_3$ : C, 49.0; H, 4.1; N, 14.3%.

(ii) *Action of Nitric Acid on the Tetrahydro-Acid (III).*—1-Methyl-5,6,7,8-tetrahydro-4-quinolone-3-carboxylic acid was refluxed for 14 hours with 68% nitric acid. The solution was evaporated to a small volume, diluted with water, made alkaline with sodium bicarbonate, and filtered. Acidification of the filtrate gave a crystalline precipitate (V) which crystallized from acetic acid as colourless leaflets, m.p. 298–299 °C. (uncorr.), mixed m.p. with the acid (II) c. 260 °C. Yield, 10–25%.

Found: C, 49.0; H, 3.8; N, 7.0;  $CH_3N$ , 14.7%.

Calculated for  $C_8H_7O_5N$ : C, 48.7; H, 3.6; N, 7.1;  $CH_3N$ , 14.7%.

No base could be isolated from the attempted decarboxylation of this acid.

The bicarbonate-insoluble solid was extracted with boiling water, which removed a little unchanged starting material (m.p. and mixed m.p.), then crystallized from acetic acid. 6-Nitro-1-methyl-4-quinolone-3-carboxylic acid (IV) separated as small cream plates, m.p. 259–261 °C (uncorr.) alone or mixed with an authentic specimen (see footnote p. 273). Yield, 50–60%.

Found: C, 53.2; H, 3.4; N, 11.2;  $CH_3N$ , 10.6%.

Calculated for  $C_{11}H_8O_5N_2$ : C, 53.2; H, 3.2; N, 11.3;  $CH_3N$ , 11.7%.

IV is a very weak acid, insoluble in bicarbonate. Its sodium salt is only sparingly soluble in the presence of excess alkali.

#### (f) 6-Nitro-1-Methyl-4-Quinolone (VI)

Decarboxylation of the acid (IV) in boiling dibutyl phthalate containing a little copper-bronze gave the weakly basic VI, short yellow needles from acetic acid, which became opaque on drying, m.p. 238–239 °C.

Found: C, 58.8; H, 3.8; N, 13.8;  $CH_3N$ , 13.0%.

Calculated for  $C_{10}H_8O_3N_2$ : C, 58.8; H, 3.9; N, 13.7;  $CH_3N$ , 14.2%.

On refluxing for seven hours with 68% nitric acid, VI was converted to 3,6-dinitro-1-methyl-4-quinolone (VIII), m.p. and mixed m.p. 269–271 °C. (uncorr.).

#### (g) 3-Amino-1-Methyl-4-Quinolone

3-Nitro-1-methyl-4-quinolone (VII) was reduced by tin and hydrochloric acid and the product extracted by chloroform after basification. Evaporation of the chloroform left a crystalline residue which was not further examined, but was converted to the picrate, dull yellow felted needles from aqueous ethanol, m.p. 234–235 °C. (decomp.).

Found: C, 47.8; H, 3.4; N, 17.1%.

Calculated for  $C_{10}H_{10}ON_2.C_6H_5O_7N_3$ : C, 47.6; H, 3.2; N, 17.4%.

The same substance was obtained in small yield (c. 5%) when the 3-nitro- compound (VII) was refluxed in benzene with excess anhydrous aluminium bromide for half to one hour. It was isolated as the picrate, m.p. 231–232 °C. (decomp.), mixed m.p. with the picrate described above 231–232 °C., mixed m.p. with the picrate of 1-methyl-4-quinolone, c. 215 °C.

#### (h) Action of Nitric Acid on 1-Methyl-4-Quinolone-3-Carboxylic Acid

The acid (II) was refluxed with 68% nitric acid for 15 hours. The solution was then evaporated to dryness on the water-bath, and the residue extracted twice with warm sodium carbonate solution. The soluble and insoluble fractions were examined as follows:

(i) The material insoluble in sodium carbonate was extracted with small portions of boiling water until the extracts, initially yellow, were colourless. The water-insoluble cream solid was extracted three times with warm 10% caustic soda, and the yellow alkaline extracts treated at once with sodium bicarbonate. The bicarbonate precipitated 3,6-dinitro-1-methyl-4-quinolone (VIII), m.p. and mixed m.p. 270–272 °C. (uncorr.). The substance insoluble in caustic soda, after crystallization first from acetic acid, then from methanol, was obtained as colourless needles, m.p. 226–228 °C., mixed m.p. with 3-nitro-1-methyl-4-quinolone (VII), 227–229 °C.

The hot water extracts were acidified, the precipitate filtered, and the filtrate discarded. The precipitate was extracted with warm 10% caustic soda, and the extract acidified. Repeated crystallization of the precipitate from acetic acid gave some impure starting material (II), m.p. 286–291 °C. (uncorr.). The caustic soda-insoluble material was dissolved in boiling water and diluted with an equal volume of 10% caustic soda. The resulting precipitate was redissolved in water and the solution acidified, giving a product which, after crystallization from acetic acid, melted at 259–261 °C. (uncorr.), not depressed by admixture with 6-nitro-1-methyl-4-quinolone-3-carboxylic acid (IV).

(ii) The sodium carbonate solutions on standing deposited a small amount of solid. This was twice dissolved in water and reprecipitated by caustic soda. Finally it was dissolved in water, the solution acidified, and the precipitate crystallized from acetic acid, giving a little more of the acid (IV), m.p. and mixed m.p. 259–261 °C. (uncorr.). The sodium carbonate filtrate on acidification precipitated some impure starting material (II), m.p. and mixed m.p. 291–293 °C. (uncorr.) after repeated recrystallization. The aqueous acid filtrates from the crude II were extracted continuously with chloroform for 12 hours. The chloroform was evaporated and the residue treated with methanol. The small insoluble fraction crystallized from acetic acid as colourless needles, m.p. 295–296 °C. (uncorr.), alone or mixed with the acid,  $C_8H_7O_3N$  (V). The mixed m.p. with II was depressed 40 °C. The deep yellow methanol solution was evaporated, the residue extracted with benzene, and the benzene-soluble material crystallized from benzene-light petroleum. Almost colourless plates of picric acid separated, m.p. and mixed m.p. 121–122 °C.

Found: C, 31.6; H, 1.4%.

Calculated for  $C_8H_3O_7N_3$ : C, 31.4; H, 1.3%.

(i) *Reduction of 1-Methyl-4-Quinolone-3-Carboxylic Acid and its Tetrahydro- Derivative with Zinc and Hydrochloric Acid*

(i) 1-Methyl-4-quinolone-3-carboxylic acid (II; 2 g.) dissolved in concentrated hydrochloric acid (50 ml.), was reduced by the gradual addition of a large excess of zinc dust. The solution was evaporated to about half the volume, poured into excess 40% caustic soda, and extracted with ether. Evaporation of the dried ethereal solution left an oily residue (1.1 g.) which crystallized from benzene-light petroleum. The base (XIII), after purification through the picrate, was obtained as colourless needles from benzene-light petroleum, m.p. 153–154 °C. The mixed m.p. with a synthetic specimen (see footnote p. 277) of 1,3-dimethyl-4-quinolone was 152.5–153.5 °C.

Found: C, 76.4; H, 6.4; N, 8.1%.

Calculated for  $C_{11}H_{11}ON$ : C, 76.3; H, 6.4; N, 8.1%.

The picrate was crystallized from methanol (charcoal) and obtained as long yellow needles, m.p. 189–190 °C., undepressed by admixture with the picrate of synthetic XIII.

Found: C, 50.9; H, 3.7; N, 14.0%.

Calculated for  $C_{11}H_{11}ON.C_6H_3O_7N_3$ : C, 50.7; H, 3.5; N, 13.9%.

(ii) The tetrahydro- acid (III; 1 g.) dissolved in hydrochloric acid was reduced with a large excess of zinc dust. There was obtained 0.8 g. crude base which was converted to the picrate. Repeated crystallization from aqueous methanol gave yellow plates, m.p. 169–171 °C.

Found: C, 50.4; H, 4.5; N, 13.5%.

Calculated for  $C_{11}H_{13}ON.C_6H_3O_7N_3$ : C, 50.2; H, 4.4; N, 13.8%.

### III. ACKNOWLEDGMENTS

This work was carried out as part of the research programme of the Division of Industrial Chemistry, C.S.I.R.O. The author is indebted to Dr. F. N. Lahey for synthetic specimens of I, II, IV, and XIII, and to Mr. R. D. Brown for a molecular weight determination.

## IV. REFERENCES

- (1) CROW, W. D.—*Aust. J. Sci. Res. A* **2**: 264 (1949).
- (2) GRESHOFF, M.—*Rec. Trav. chim. Pays-Bas* **19**: 360 (1900).
- (3) SPÄTH, E., and KOLBE, A.—*Mh. Chem.* **43**: 469 (1923).
- (4) EHRENSTEIN, M., and BUNGE, W.—*Ber. deutsch. chem. Ges.* **67**: 1715 (1934).
- (5) VON NIEMENTOWSKI, S., and SUCHARDA, E.—*J. prakt. Chem.* **94**: 193 (1916).
- (6) PAPA, D., SCHWENK, E., and WHITMAN, B.—*J. Org. Chem.* **7**: 587 (1942).
- (7) TAYLOR, T. W. J., and BAKER, W.—“*Sidgwick's Organic Chemistry of Nitrogen.*” p. 259.  
(Clarendon Press: Oxford, 1942.)
- (8) SEYFFERTH, E.—*J. prakt. Chem.* **34**: 241 (1886).
- (9) BAEYER, A.—*Liebigs Ann.* **245**: 103 (1888).
- (10) CROW, W. D., and PRICE, J. R.—*Aust. J. Sci. Res. A* **2**: 255 (1949).

# ALKALOIDS OF THE AUSTRALIAN RUTACEAE: *MELICOPES* *FAREANA*

## V. THE STRUCTURE OF THE ALKALOIDS

By W. D. CROW\* and J. R. PRICE\*

[Manuscript received March 9, 1949]

### Summary

Melicopine, melicopidine, and melicopine are shown to be members of a new group of alkaloids derived from acridine. The structure of melicopine, 1,2,3,4-tetramethoxy-10-methylacridone (II), is deduced from data reported in earlier papers. The presence of the same 10-methylacridone skeleton in melicopidine and melicopine is established by conversion of the trimethoxyphenols obtained from them by fission of the methylenedioxy ring with methanolic potash to the respective dimethoxy-*o*- and *p*-quinones previously prepared from melicopine. This conversion also establishes the position of the methylenedioxy group in melicopine relative to the hydroxyl group in *normelicopine*. Similar considerations applied to the ethoxydimethoxyphenols show the position of the methylenedioxy group in melicopidine relative to the hydroxyl group in *normelicopine*, and in this case, lead to the complete structure for the alkaloid (XIII).

The action of nitrous acid on *normelicopine* and *normelicopidine* gives two hydroxymethoxyquinones isomeric with that obtained by the action of sodium carbonate on the dimethoxy-*o*- and *p*-quinones. The same two hydroxymethoxyquinones also result from the action of caustic soda on the dimethoxy-*o*- and *p*-quinones respectively. Their structures, which can be deduced from the second method of preparation, confirm the positions of the methylenedioxy group in melicopine and melicopidine relative to the hydroxyl group of *normelicopine*, and prove that the hydroxyl groups in *normelicopine* and *normelicopidine* are in the same position as that in *normelicopine*, but they do not make possible a choice between alternative structures for melicopine. This choice depends on the position of the hydroxyl group of the *noralkaloids* relative to the remainder of the acridone molecule. By consideration of the mechanism of fission of the methylenedioxy ring, the hydroxyl group of the *noralkaloids* is shown to be situated *peri* to the acridone oxygen atom, i.e. at position 4. This is confirmed by the occurrence of hydrogen bonding. Consequently the complete structural formula for melicopine is XXIII.

The properties of the alkaloids and a number of the degradation products are discussed.

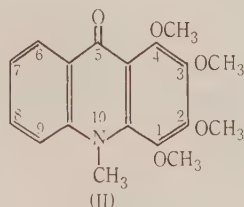
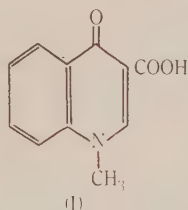
### I. INTRODUCTION

The formation of 1-methyl-4-quinolone-3-carboxylic acid (I) from melicopine, melicopidine, and melicopine(I) implies that the alkaloids are related to the quinolone by fusion of a third ring in the 2,3 positions. It is improbable that any rearrangement of the original skeleton has taken place in the formation of this acid as it can be obtained under mild alkaline conditions from the dihydroxy-

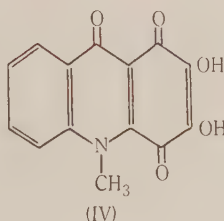
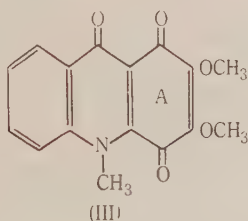
\* Division of Industrial Chemistry, C.S.I.R.O.



quinone,  $C_{14}H_9O_5N$ , in which the original structure is known to be present(2), as well as by the vigorous acid treatment described in Part IV. In the degradation of melicopicine,  $C_{18}H_{19}O_5N$ , to 1-methyl-4-quinolone,  $C_{10}H_9ON$ , a fragment,  $C_8H_{10}O_4$ , is lost. This fragment includes four methoxyl groups and can only be accommodated by fusion of a tetramethoxybenzenoid ring to the quinolone nucleus. That is, melicopicine must be 1,2,3,4-tetramethoxy-10-methylacridone (II).



The possibility referred to in Part III(2) that one of the dimethoxyquinones obtained from melicopicine may be a bicyclic quinone rather than a *p*-quinone is now excluded; this substance will have structure III and the dihydroxyquinone structure IV.



The dimethoxy-*p*-quinone (III) may also be obtained from *normelicopicine*; consequently the methoxyl group demethylated in the formation of *normelicopicine* must be situated at either the 1 or the 4 position. Until a decision can be made between these alternatives, there will also be two possible structures for the dimethoxy-*o*-quinone, the hydroxymethoxyquinone(2), and a number of related substances to be described in this paper. In the meantime it will be sufficient to consider the arrangement of substituents in ring A only, without relation to the remainder of the acridone molecule. For this purpose an arbitrary system of numbering, in which the hydroxyl group in *normelicopicine* is at position 4, will be employed for the ring A partial formulae. Such partial structures can be deduced for melicopine and melicopidine in two ways: (i) from the relation between the dimethoxyquinones obtained from melicopicine and the hydroxytrialkoxo compounds obtained from melicopine and melicopidine, and (ii) from the relation between the isomeric hydroxymethoxyquinones derived from the three alkaloids.

(a) *The Action of Nitric Acid on the Hydroxyalkoxy-Compounds obtained from Melicopine and Melicopidine*

It has been shown(3) that methanolic potash replaces the methylenedioxy group in melicopine and melicopidine by a hydroxyl and a methoxyl group, giving in each case a trimethoxyphenol. Both these phenols are isomeric with *normelicopicine*. That from melicopine is converted by nitric acid to a quinone identical with the 2,3-dimethoxy-1,4-quinone (III). Further, the phenol is demethylated by alcoholic acid to a dihydroxydimethoxy compound(3) which proves to be the quinol corresponding to III. It follows that melicopine is derived from 1,2,3,4-tetrahydroxy-10-methylacridone.

The trimethoxyphenol resulting from fission of the methylenedioxy ring of melicopidine is likewise converted by nitric acid to a quinone, in this case

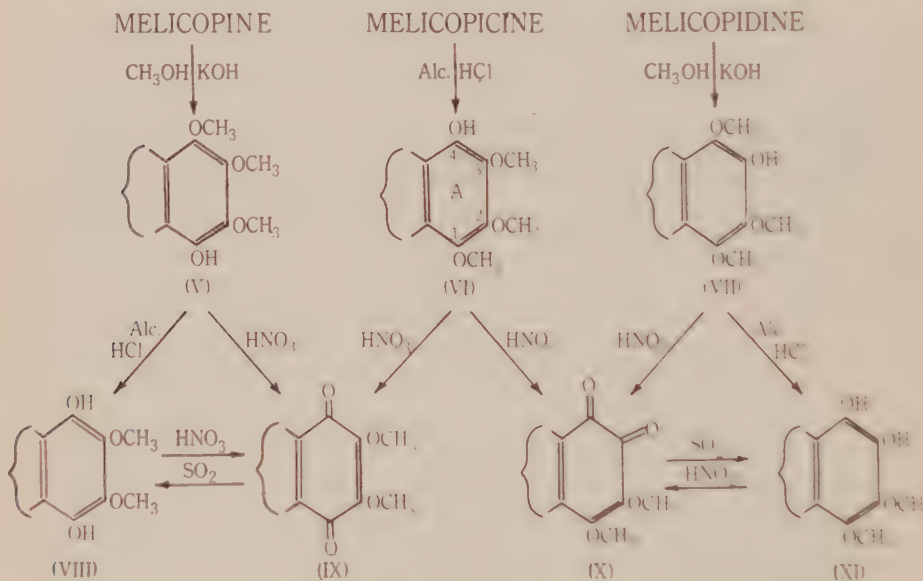
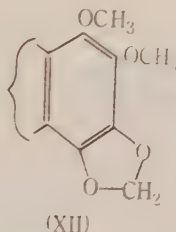


Fig. 1

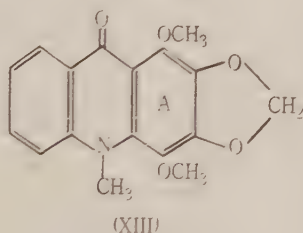
identical with the 1,2-dimethoxy-3,4-quinone from melicopicine. The corresponding quinol results from demethylation of the phenol with alcoholic acid. Consequently, melicopidine also is derived from 1,2,3,4-tetrahydroxy-10-methylacridone.

Since the same dimethoxy-1,4-quinone (III) can be obtained either from *normelicopicine* or from the trimethoxyphenol from melicopine, and since the former is assigned the 4-hydroxy-1,2,3-trimethoxy structure (VI), it follows that the latter must be the corresponding 1-hydroxy-2,3,4-trimethoxy-compound (V). Similarly, the trimethoxyphenol from melicopidine, since it yields the dimethoxy-3,4-quinone, also obtained from *normelicopicine*, must be the 3-hydroxy-1,2,4-trimethoxy-compound (VII). Moreover, the methoxyl group demethylated by alcoholic acid must occupy the same position in both trimethoxyphenols (V and VII) as in melicopine. These relationships between the three alkaloids are shown in Figure 1.

As the hydroxyl group in the trimethoxyphenol (V) arises by fission of a methylenedioxy ring, this methylenedioxy ring must be at the 1,2 positions in melicopine, which is therefore represented by the partial structure XII.



With ethanolic potash, melicopine and melicopidine give isomeric ethoxydimethoxyphenols(3). That from melicopine is converted to a *p*-quinone by nitric acid and is demethylated to the corresponding *p*-quinol by alcoholic hydrochloric acid. The quinol and the quinone both contain one methoxyl and one ethoxyl group.\* By similar changes an *o*-quinone and an *o*-quinol can be obtained from melicopidine, each containing one methoxyl and one ethoxyl group. The quinone from melicopidine forms a *phenazine* on treatment with *o*-phenylenediamine. Just as the dimethoxy-*o*- and *p*-quinones are both hydrolysed by aqueous sodium carbonate to the same hydroxymethoxyquinone,  $C_{15}H_{11}O_5N(2)$ , so the two methoxyethoxyquinones are hydrolysed by sodium carbonate to the same hydroxyalkoxyquinone,  $C_{16}H_{13}O_5N$ . This contains an ethoxyl group, since with hydrobromic acid it loses  $C_2H_4$  giving the dihydroxyquinone (IV). The ethoxyl group occupies the same position in the *o*-quinone derived from melicopidine as in the *p*-quinone derived from melicopine, and must be in the 2 position, as it arises in each case from the fission of a methylenedioxy ring which, in melicopine is in the 1,2 position. Consequently the methylenedioxy group of melicopidine must be situated 2,3. Ring A being symmetrically substituted, the complete structural formula (XIII) can be written for melicopidine.



\* Since the trimethoxyphenol and the ethoxydimethoxyphenol, obtained from melicopine by methanolic and ethanolic potash respectively, both give *p*-quinols on treatment with alcoholic acid, it follows that it is the same methoxyl group which is demethylated in each case.



(b) *The Structures of the Isomeric Hydroxymethoxyquinones*

It has been shown in Part III(2) that *normelicopine* (VI) is converted to a mixture of the dimethoxy-*o*- and *p*-quinones (X and IX) by nitrous acid. *nor*Melicopine and *normelicopidine* are also attacked by nitrous acid (1 mole) undergoing demethylenation and oxidation to hydroxymethoxyquinones isomeric with that produced by the action of aqueous sodium carbonate on the dimethoxy-*o*- and *p*-quinones. The product from *normelicopine* is a yellow to greenish-yellow acidic substance which is unstable, cannot be recrystallized, and consequently has not been obtained sufficiently pure for analysis.\* However, its derivatives are stable and can be purified. It gives a bright red *monoacetate*,  $C_{15}H_{10}O_4N(OOC.CH_3)$ , and a violet-red *phenazine*,  $C_{21}H_{15}O_3N_3$ , and is reduced by sulphur dioxide to an orange-red *phenol*,  $C_{15}H_{13}O_5N$ , containing one methoxyl but no methylenedioxy group. This phenol which is rapidly reoxidized in alkaline solution by air to the greenish-yellow acidic substance, on acetylation gives a *diacetate*, also obtained by reductive acetylation of the parent substance. Further acetylation of the diacetate gives a *triacetate*. Hence the acid is a hydroxymethoxyquinone,  $C_{15}H_{11}O_5N$ . It is demethylated to the dihydroxyquinone (IV) by 46 per cent. hydrobromic acid, and yields melicopine on reductive methylation.

A third hydroxymethoxyquinone is formed from *normelicopidine* by treatment with nitrous acid. This quinone is even less stable than that from *normelicopine*. Its nature is established by reduction to a *phenol*,  $C_{15}H_{13}O_5N$ , which gives a *diacetate* and a *triacetate*, and which can be methylated to melicopine. The quinone itself can be demethylated to the dihydroxyquinone (IV). It reacts with acetic anhydride and with *o*-phenylenediamine, but in each case a mixture of products results.

Although the dimethoxy-*o*- and *p*-quinones (IX and X) are both hydrolysed by hot sodium carbonate solution to the same hydroxymethoxyquinone, which must therefore be the 3-hydroxy-2-methoxy- compound (XV, Fig. 2), they behave differently with caustic soda. The dimethoxy-*p*-quinone (IX) gives that hydroxymethoxyquinone which is obtained from *normelicopine* by treatment with nitrous acid. This must therefore be the 2-hydroxy-3-methoxy-1,4-quinone (XIV). It is also obtained by the action of caustic soda on the methoxyethoxy-*p*-quinone from melicopine which is therefore the 3-methoxy-2-ethoxy-1,4-quinone (XIX). Treatment of the dimethoxy-*o*-quinone (X) or of the methoxyethoxy-*o*-quinone from melicopidine with caustic soda, results in the formation of the third hydroxymethoxyquinone—that produced by the action of nitrous acid on *normelicopidine*. It follows that this is the 2-hydroxy-4-methoxy-3,4-quinone (XVI) and that the methoxyethoxy-*o*-quinone is the 1-methoxy-2-ethoxy-3,4-quinone (XXI).

The formation of the 2-hydroxy-3-methoxy-1,4-quinone (XIV) from *normelicopine* confirms that the methylenedioxy group in melicopine is in the 1,2 position. Moreover, in melicopine, the methoxyl group demethylated by

\* Analysis of the crude material, dried in a vacuum at room temperature, is in agreement with the formula  $C_{15}H_{11}O_5N.H_2O$ .



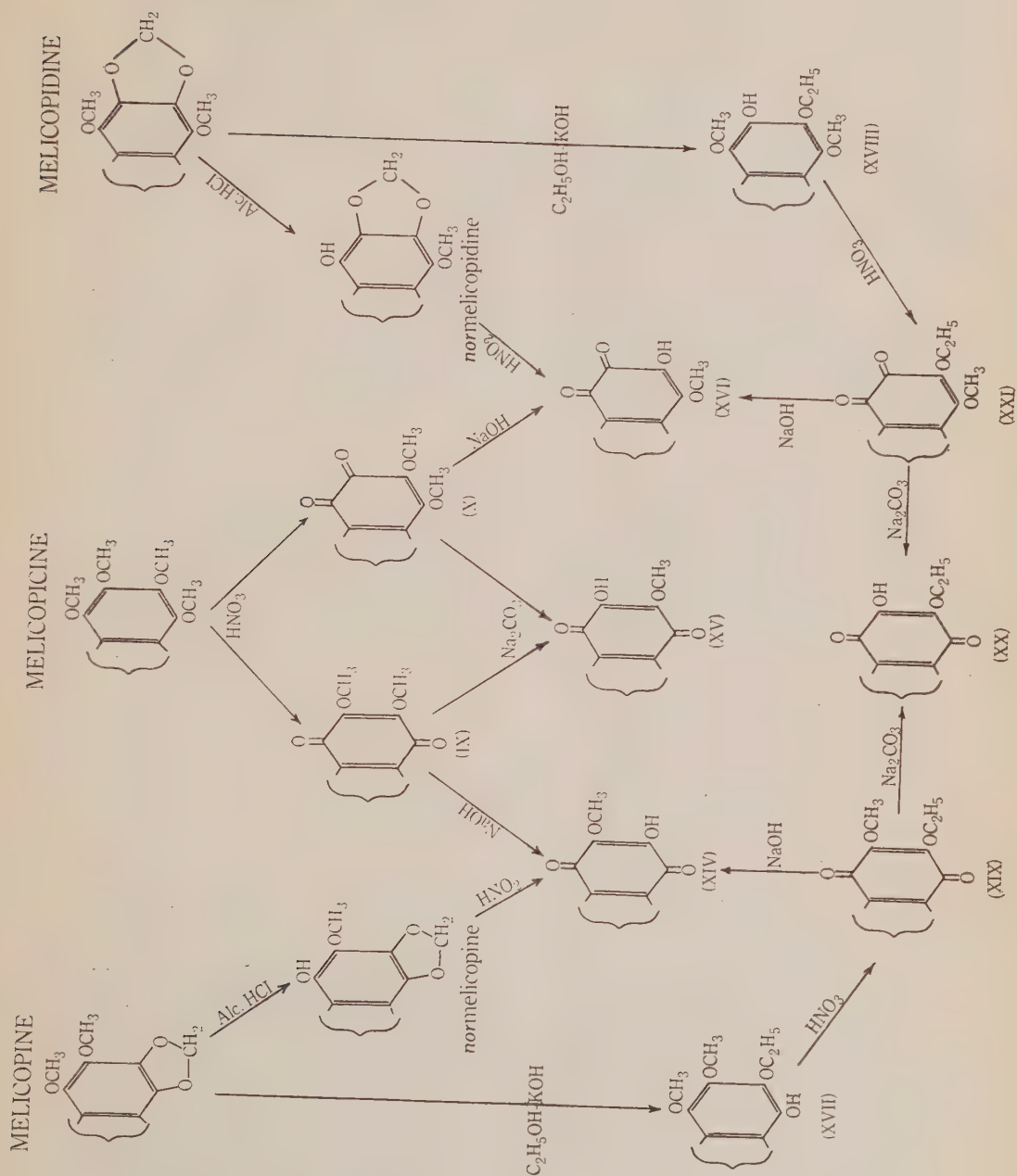


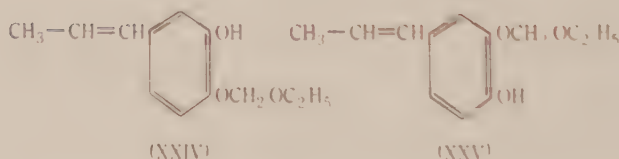
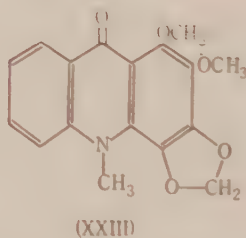
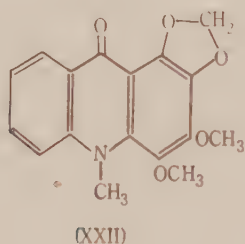
Fig. 2

alcoholic acid must be in position 4, as in melicopine. The formation from *normelicopidine* of the 2-hydroxy-1-methoxy-3,4-quinone (XVI), taken in conjunction with the previously deduced 2,3 position for the methylenedioxy group, shows that, in melicopidine also, the methoxyl group demethylated by alcoholic acid is in position 4. In other words, as would be expected, it is always the same methoxyl group which is demethylated by alcoholic acids.

However, since the numbering employed in the preceding discussion does not relate to the acridone molecule as a whole, there still remain two possible structures for the *noralkaloids* and consequently for the majority of the degradation products referred to above.

### (c) The Structure of *Melicopine*

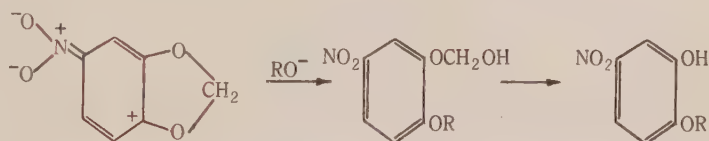
Complete structural formulae can be written for melicopine (II) and melicopidine (XIII), but there remain two possible structures (XXII and XXIII) for melicopine. A decision between the two can be made by consideration of the mechanism involved in the opening of the methylenedioxy ring by alcoholic potash.



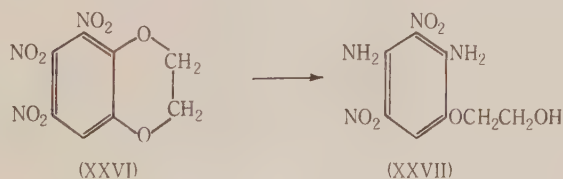
Fission of a methylenedioxy ring by alcoholic alkali or sodium alkoxides may follow one of two courses. The first, giving rise to alkoxymethyl ethers, is exemplified by *isosafrole* (or *safrole*) which gives a mixture of the isomeric 4-propenylcatechol ethoxymethyl ethers (XXIV and XXV) when heated with sodium ethoxide or alcoholic potash at 140–170 °C. (4, 5).

The second type of fission results in the replacement of the methylenedioxy group by a hydroxyl and an alkoxyl group. Thus 4-nitro-1,2-methylenedioxybenzene reacts with sodium methoxide in methanol to give 5-nitro-2-methoxyphenol, and with sodium ethoxide in ethanol to give 5-nitro-2-ethoxyphenol (6). Similarly 4,5-dinitro-1,2-methylenedioxybenzene is converted to 4,5-dinitro-2-methoxyphenol with sodium methoxide in methanol and to the corresponding

ethyl ether with sodium ethoxide in ethanol(7). Robinson and Robinson(6) have pointed out the resemblance of this type of fission to ether interchange in such substances as 2,4-dinitroanisole. Ether interchange reactions of the same kind are met with in the furanoquinoline alkaloids, skimmianine and  $\gamma$ -fagarine(8), 4-nitro-1-naphthyl ethers(9) and 6-methoxy-7-nitrobenzothiazole(10). The first step in the action of alcoholic alkoxides on 4-nitro-1,2-methylenedioxybenzene can be formulated as a straightforward nucleophilic displacement. This would be followed by immediate hydrolysis of the resulting hemiacetal\*—compare the conversion of methylenedioxybenzenes to monohydric phenols by alkaline reducing agents such as sodium amalgam(13), Raney alloy and caustic soda(14), and sodium in liquid ammonia(15).



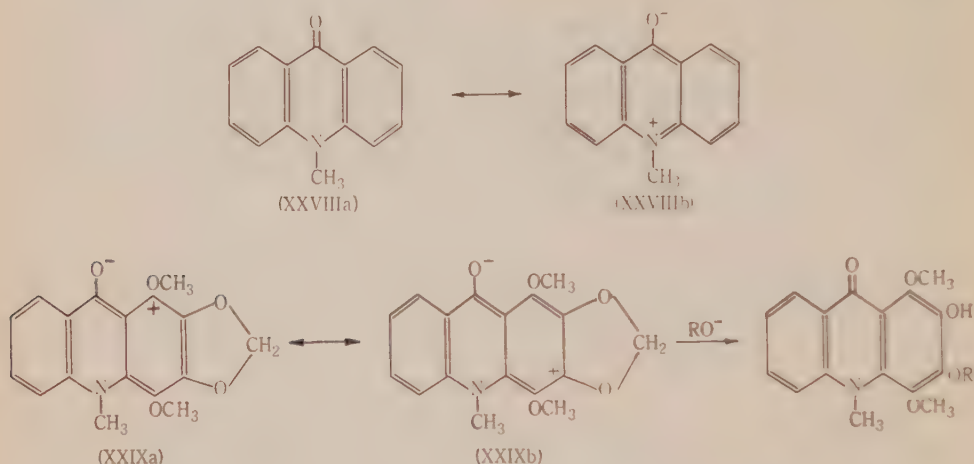
The acetal structure of the methylenedioxy ring evidently facilitates fission as the reaction does not take place with 4-nitro-1,2-ethylenedioxybenzene. However, nucleophilic attack is promoted by increasing the number of electron attracting substituents; ammonia in aqueous pyridine converts 3,4,5-trinitro-1,2-ethylenedioxybenzene (XXVI) to 2,4-diamino-3,5-dinitro- $\beta$ -hydroxyethoxybenzene (XXVII)(6). In this instance ring opening leads to a glycol monoether, which would not be expected to hydrolyse further.



Of the two kinds of ring fission which could be brought about by alcoholic alkali, the behaviour of melicopine and melicopidine is clearly analogous to that of 4-nitro-1,2-methylenedioxybenzene. This being so, application of the ether interchange mechanism should make possible a decision between the alternative positions which could be taken up by the entering alkoxy group, and consequently a decision as to the location of the hydroxyl group of the *nor*alkaloids. All three alkaloids are weak bases owing to resonance between structures such as XXVIIIa and XXVIIIb(16). Minor contributions from quinonoid structures are also to be expected, for example XXIXa and XXIXb; these would assume some importance on attack by a nucleophilic reagent such as the alkoxy ion, and ether interchange should only occur at position 2 or 4.

\* Some interesting examples of this type of ring fission are reported by Slooff(11) and Böeseken and Slooff(12).

With melicopidine, attack must take place at the 2 position (since the 4 position cannot be involved in the methylenedioxy ring) resulting, as shown below, in the formation of a 2-alkoxy-3-hydroxy-1,4-dimethoxy-10-methylacridone.

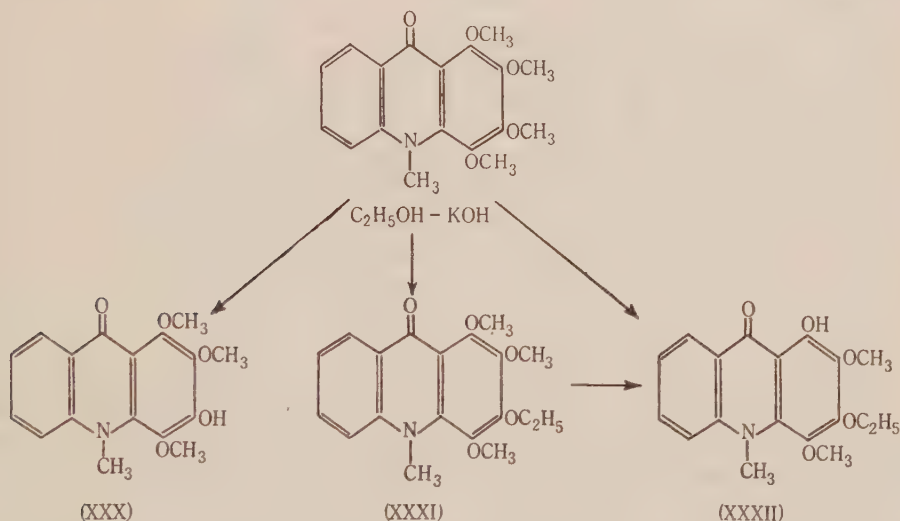


Since the ethoxydimethoxyphenol (XVIII) obtained from melicopidine and ethanolic potash is converted by alcoholic hydrochloric acid to an *o*-quinol which still contains the ethoxyl group, the methoxyl group demethylated by alcoholic acid must be at position 4. Further, the ethoxyl group has been shown to be in the same position in the ethoxydimethoxyphenol from melicopidine as in the corresponding ethoxydimethoxyphenol (XVII) from melicopine. It follows that the methylenedioxy group in melicopine must be in the 1,2 position, so that this alkaloid is correctly represented by XXIII. Three observations substantiate the mechanism of fission of the methylenedioxy ring and consequently the 1,2 - methylenedioxy - 3,4 - dimethoxy - 10 - methylacridone structure for melicopine.

- (1) The general properties of the *nor*alkaloids are in complete harmony with the allocation of the hydroxyl group to position 4 (*vide infra*).
- (2) Melicopine reacts slowly with ethanolic potash giving a mixture of phenolic and non-phenolic products. The phenol,  $C_{17}H_{17}O_5N$ , gives a *monoacetate* and can be methylated to melicopine. It is not identical with any of the three monohydroxytrimethoxy- compounds described previously and must therefore be 2-hydroxy-1,3,4-trimethoxy-10-methylacridone (XXX). It is of interest in that it resembles *normelicopine* in not being further demethylated by alcoholic acid. The non-phenolic products are separable into two fractions. The more weakly basic fraction consists of a substance,  $C_{18}H_{19}O_5N$ , isomeric with melicopine but resembling *normelicopine* in its properties. It is not further demethylated by alcoholic acid, but is converted by nitric acid to the 2-ethoxy-3-methoxy-1,4-quinone (XIX). Since it is not the 1-hydroxy-2-ethoxy-3,4-dimethoxy- compound it must be 4-hydroxy-



2-ethoxy-1,3-dimethoxy-10-methylacridone (XXXII). The less weakly basic fraction gives a crystalline hydrochloride but the base could not be obtained in a solid state. However, it is converted by alcoholic acid to XXXII and so must be the 4-methyl ether (XXXI), or possibly the ethyl ether, of XXXII. Though the phenol (XXX) results only from demethylation, ether interchange at position 2 has taken place in the formation of XXXI and XXXII. This is to be expected as the reaction with melicopine and melicopidine is not primarily dependent on the presence of a methylenedioxy group.

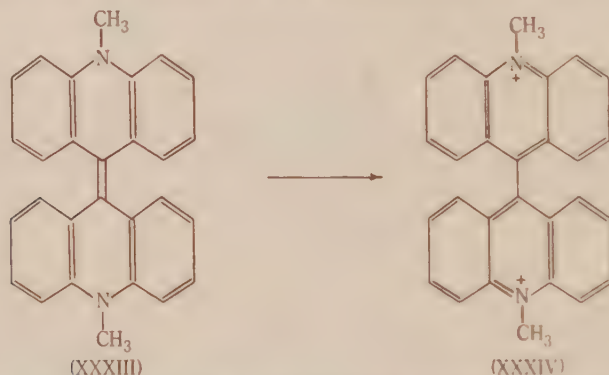


- (3) Activation at position 4 (cf. structure XXIXa) could also lead to ether interchange, and there is reason to suppose that with melicopine, interchange does take place with the 4-methoxyl group to a small extent. The small non-phenolic fraction remaining after treatment of melicopine with ethanolic potash is a mixture of melicopine and *normelicopine* with a yellow crystalline base resembling melicopine but melting  $30^\circ C.$  lower. It cannot be separated into components by fractional crystallization or by differences in basic strength but appears to be a mixture (or loose addition compound) of melicopine with a similar substance. Treatment with alcoholic acid gives *normelicopine* (89 per cent. of the weight of starting material). It follows that the second substance can only differ from melicopine in the nature of the substituent at the 4 position, and is probably formed by exchange of the methoxyl at this position for ethoxyl.

The location of the hydroxyl group of the *noralkaloids* enables complete structural formulae to be written for the various phenols, quinones, and hydroxyquinones represented previously by partial formulae (V-XI, XIV-XXI).

(d) *Properties of the Alkaloids and their Degradation Products*

On the evidence submitted above, melicopine, melicopidine, and melicopicine are substituted *N*-methylacridones and thus constitute the first recorded occurrence of acridine derivatives in the plant kingdom. Additional confirmation of the acridone structure is supplied by the ultraviolet absorption spectra, which have been studied by Brown and Lahey (unpublished data), and by the reduction to diacridenes. 10-Methylacridones are reduced by zinc dust in acetic (or hydrochloric) acid to 10,10'-dimethyldiacridenes (XXXIII) which with dilute nitric acid can be oxidized to the diacridinium ions (XXXIV)(17).



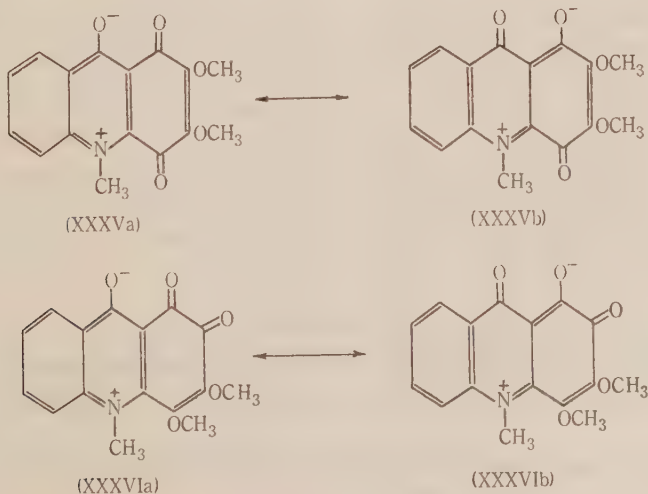
Melicopine, melicopidine, and melicopicine are all reduced by zinc dust and hydrochloric acid, but only the product from melicopine, 1,2,1',2'-bismethylenedioxy-3,4,3',4'-tetramethoxy-10,10'-dimethyldiacridene, has been prepared. It is a yellow fluorescent solid,  $C_{34}H_{30}O_8N_2$ , and is rapidly oxidized by hot dilute nitric acid to a permanganate-coloured diacridinium dinitrate, which has been converted to the *dipicrate*,  $C_{34}H_{30}O_8N_4 \cdot 2C_6H_2O_7N_3$ .

In view of the very close structural similarity of the three alkaloids, it is difficult to account for their different basicities. Melicopidine is the strongest base and melicopicine the weakest, though not much weaker than melicopine. That is, replacement of two methoxyl groups in melicopicine by methylenedioxy increases the basicity appreciably in the 2,3 positions but very little in the 1,2 positions. Solutions of melicopidine in aqueous acid show unusual behaviour in that the salt separates on dilution. It is possible that the increased solubility at lower pH is due to oxonium salt formation, which might also account for the differences between the three alkaloids.

Of the four hydroxytrimethoxy-10-methylacridones derived from the alkaloids, the 4-hydroxy- compound, *normelicopicine*, differs in many respects from the isomeric 1-, 2-, and 3-hydroxy compounds. These have normal phenolic properties and resemble the alkaloids in colour, solubility, and basicity. *nor*Melicopicine, like *normelicopine* and *normelicopidine*, is insoluble in alkali and is an extremely weak base. Further, the *nor*alkaloids are more highly coloured, more soluble in non-polar solvents and less soluble in polar solvents than the 1-, 2-, or 3-hydroxy compounds. These differences can be ascribed to hydrogen bond formation between the 4-hydroxyl group of the *nor*alkaloids

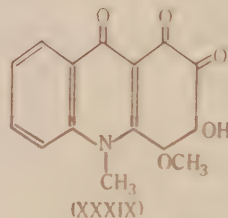
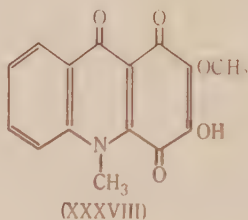
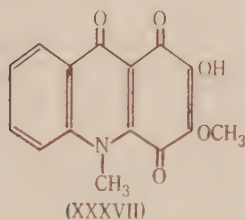
and the acridone oxygen atom, which also accords with the observation that the wet melting point of *normelicopicine* is only 7–10 °C. below the true melting point, whereas there is a depression of *c.* 45 °C. with the 3-hydroxy- isomer. The point is established by comparison of the infra-red absorption spectra of the four monohydroxytrimethoxy-*N*-methylacridones, which were examined in the solid state. In addition to the characteristic C–H absorption bands near 2900  $\text{cm}^{-1}$ , the 1-, 2-, and 3-hydroxy- compounds show an absorption band at 3279, 3279, and 3254  $\text{cm}^{-1}$  respectively, which is attributable to a “free” O–H group. It is not present in the spectrum of *normelicopicine* in which the hydroxyl group must therefore be situated *peri* to a carbonyl group. The low values of the O–H bands in the 1-, 2-, and 3-hydroxy compounds are ascribed to intermolecular association. Since hydrogen bond formation will be favoured by an increased contribution from zwitterionic structures (of type XXXVIIIb), the basicity of the *nor*alkaloids should be lower than that of the alkaloids. The ease with which demethylation takes place at the 4 position is presumably associated with the high degree of stability conferred on the products by hydrogen bonding. Oxidative dealkylation of the three *nor*alkaloids to quinones occurs with surprising facility with one mole of nitrous acid. The reaction resembles the conversion by nitrous acid of quinol monomethylether to benzoquinone(18). Although some 1,2-dimethoxy-3,4-quinone as well as the isomeric 1,4-quinone is formed from *normelicopicine*, only one product could be isolated from *normelicopine* and *normelicopidine*. Evidently the methylenedioxy group is attacked preferentially.

In both dimethoxyquinones the influence of the keto- group at position 4 is probably diminished by resonance between structures XXXVa and XXXVb or XXXVIa and XXXVIb. Consequently demethylation should be controlled by the keto- group at position 1 in the *p*-quinone and by that at position 3 in the *o*-quinone. In other words the *p*-quinone should undergo demethylation at the 3 position and the *o*-quinone at the 1 position, as in fact they do with aqueous sodium carbonate, both giving the 3-hydroxy-2-methoxy-1,4-quinone (XXXVII).





However, no explanation can be offered for the remarkable behaviour of both dimethoxyquinones with caustic soda, which brings about demethylation of the 2-methoxyl group. Moreover, the properties of the products of caustic soda hydrolysis, the 2-hydroxy-3-methoxy-1,4-quinone (XXXVIII), and the 2-hydroxy-1-methoxy-3,4-quinone (XXXIX), are anomalous. Whereas the dimethoxyquinones and the 3-hydroxy-2-methoxyquinone (XXXVII) are orange-red to red in colour (in the solid state and in solutions) the 2-hydroxy-3-methoxyquinone (XXXVIII) is a greenish-yellow unstable substance whose solutions in chloroform and other non-hydroxylic solvents are green.



Reduction of XXXVIII gives the orange-red quinol resembling in colour the quinol from XXXVII. That is, in so far as the colour of the solid quinone is concerned, reduction has a bathychromic instead of the usual hypsochromic effect. Acetylation of XXXVIII, like reduction, brings it into line with XXXVII as the acetyl derivatives of the two substances are both bright red beautifully crystalline solids, indistinguishable except by their melting points. As shown in Part III(2), XXXVII is stable to boiling 10 per cent. caustic soda but under the same conditions XXXVIII is rapidly destroyed, the main product being a base,  $C_{13}H_{13}O_3N$ , containing one methoxyl group.\* The infra-red absorption spectra of XXXVII and XXXVIII were examined and, for comparison, the spectra of the dimethoxy-*p*-quinone (IX) and 2-hydroxy-1,4-naphthoquinone. In this respect also, XXXVIII appears to be anomalous. The C=O frequency in the spectrum of IX ( $1683\text{ cm.}^{-1}$ ) is the same as in the spectrum of 2-hydroxy-1,4-naphthoquinone which in addition shows an O-H band at  $3170\text{ cm.}^{-1}$ . The spectrum of XXXVII is in agreement with these two, though there is a slight shift of the two bands to lower frequencies (C=O,  $1669\text{ cm.}^{-1}$ ; O-H,  $3110\text{ cm.}^{-1}$ ). But with XXXVIII the shift of the C=O band ( $1628\text{ cm.}^{-1}$ ) is of the order encountered with strong intramolecular hydrogen bonding though an O-H band is still present ( $3270\text{ cm.}^{-1}$ ). Flett(19) found that hydrogen bonding in hydroxyanthraquinones resulted in disappearance of the O-H band.

### (c) Acid Degradation Products

In Parts II(3) and III(2) a number of acids were described which are obtained from oxidation experiments or as by-products in other reactions. The colourless

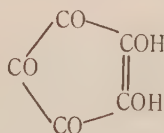
\* Presumably 1-methyl-3-methoxyacetyl-4-quinolone; it gives a 2,4-dinitrophenylhydrazone and is oxidized to I by acid permanganate.



dibasic acid,  $C_{13}H_9O_6N$ , the principal product from the aerial oxidation of alkaline solutions of 1-methoxy-2-alkoxy-3,4-dihydroxy-10-methylacridones, may be obtained more conveniently by the action of excess nitrous acid on melicopidine. The acid is converted to 1-methyl-4-quinolone by heating with zinc dust and to 1-methyl-4-quinolone-3-carboxylic acid (I) by oxidation with 68 per cent. nitric acid or with acid permanganate. Decarboxylation gives a base,  $C_{11}H_9O_2N$ , together with a little I. The base forms a 2,4-dinitrophenylhydrazone and is easily oxidized to I; it is therefore 1-methyl-4-quinoline-3-aldehyde (XLIII). This behaviour, characteristic of  $\alpha$ -keto-acids, shows the presence of a  $-CO.COOH$  substituent at position 3. Since the second carboxyl group must be at position 2, the acid,  $C_{13}H_9O_6N$ , has structure XLI.

A minor product from the aerial oxidation of the dialkoxy-*o*-quinols is a colourless acid,  $C_{13}H_9O_6N$  (m.p. 254–256 °C.). This acid may be obtained in good yield by oxidation of the 2-hydroxy-3-methoxy-1,4-quinone in caustic soda solution with permanganate. It is oxidized to I and decarboxylated to a base,  $C_{12}H_9O_3N$ , having lactonic properties. The structures of the acid (XLII) and its decarboxylation product (XLIV) are established by the formation of the former from the acid,  $C_{13}H_9O_6N$ , by reduction with zinc dust and caustic soda—compare the reduction of phthalonic acid to phthalide carboxylic acid by Scherks(20) and by Graebe and Trümpy(21).

A third substance,  $C_{13}H_9O_4N$ , is a by-product of several reactions. It can be prepared in quantity in several ways, such as by aerial oxidation of alkaline solutions of the dialkoxy-*p*-quinols or by the action of excess nitrous acid on normelicopine. It is a weakly acidic orange solid, aqueous alkaline solutions of which are yellow with a pronounced green fluorescence. The substance does not appear to contain a carboxyl group, could not be acetylated, and did not react with 2,4-dinitrophenylhydrazine. Oxidation with 68 per cent. nitric acid gives 1-methyl-4-quinolone-3-carboxylic acid, and with cold alkaline permanganate the acid,  $C_{13}H_9O_6N$  (XLI). The only structure which could account satisfactorily for these properties is XL, resembling croconic acid (XLV) which likewise fails to acetylate(22).



(XLV)

The formation of XL by the aerial oxidation of the dialkoxy-*p*-quinols does not involve the quinone as an intermediate. Starting with the *p*-quinone under the same conditions, the reaction does not proceed beyond the 2-hydroxy-3-methoxy-1,4-quinone.

The relation between the three  $C_{13}$  acids is shown in Figure 3.

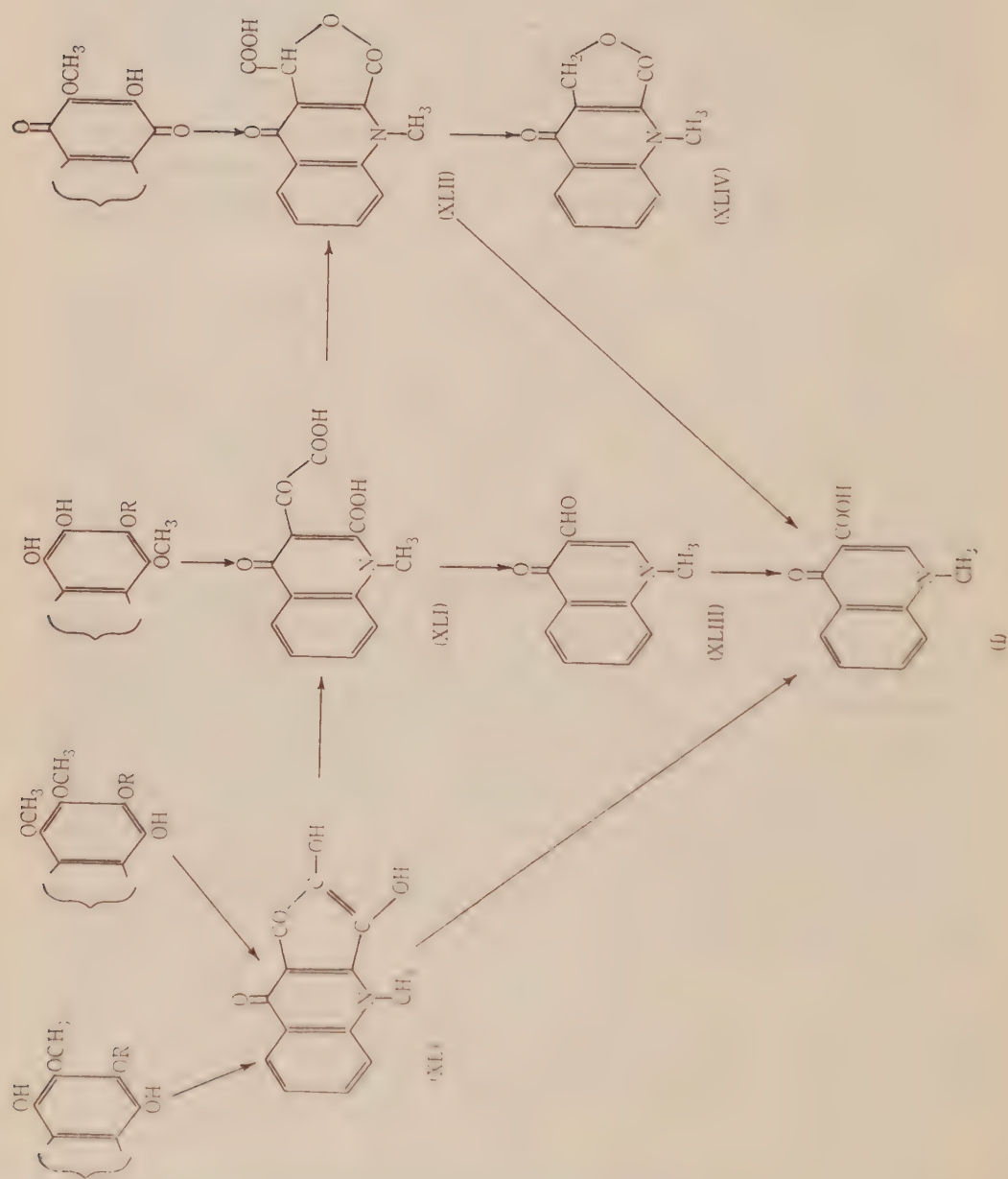
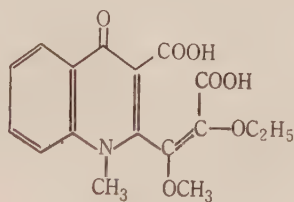


Fig. 3

Although the oxidation of 1-hydroxy-2-ethoxy-3,4-dimethoxy-10-methyl-acridone with alkaline hydrogen peroxide gives the same product (XL) as the aerial oxidation in alkaline solution of the related *p*-quinol, the 3-hydroxy-2-ethoxy-1,4-dimethoxy- compound does not give with alkaline hydrogen peroxide the same acid (XLI) as results from aerial oxidation of an alkaline solution of the related *o*-quinol. Instead, the product is a colourless dibasic acid,  $C_{17}H_{17}O_7N$ , containing two alkoxy groups. Since only one carbon atom is lost, ring fission must take place between carbon atoms 3 and 4 leading to structure XLVI for the acid, whereas in the aerial oxidation of the related *o*-quinol either the 1-2 or the 2-3 bond must be broken. The *o*-quinones cannot be intermediates in the oxidation of the dialkoxy-*o*-quinols, as none of the acid (XLI) is formed from the *o*-quinones under the same conditions. These reactions afford an illustration of the ease and specificity with which fission of highly oxygenated aromatic rings can take place. The acid,  $C_{17}H_{17}O_7N$ , also arises as a by-product in the action of ethanolic potash on melicopidine. It readily undergoes partial decarboxylation to an acid,  $C_{16}H_{17}O_5N$ .



(XLVI)

## II. EXPERIMENTAL

All melting points are corrected unless otherwise stated. Microanalyses were carried out by R. B. Bradbury and H. L. Oates.

### (a) Oxidation of the Dihydroxyquinone (IV)

The dihydroxyquinone dissolved in hot 5% caustic soda was oxidized by the gradual addition of hydrogen peroxide until the solution was almost colourless. Acidification afforded a white crystalline precipitate which, after recrystallization from acetic acid, sintered at 280 °C., and melted at 294–296 °C. (uncorr.). The mixed melting point with authentic 1-methyl-4-quinolone-3-carboxylic acid was undepressed.

### (b) Action of Nitric Acid on the Alkoxydimethoxyphenols

(i) *1-Hydroxy-2,3,4-Trimethoxy-10-Methylacridone* (V).—The trimethoxyphenol obtained from melicopine was suspended in water and 68% nitric acid added until solution was complete. When the reaction ceased (1–2 minutes) the solution was poured on to excess sodium bicarbonate, and extracted with chloroform. Removal of the solvent and crystallization of the residue from ethanol gave the dimethoxy-*p*-quinone (IX), bright red needles, m.p. 200–201 °C., mixed m.p. with a specimen prepared from melicopidine 200–201 °C. The yield was c. 90%. The identity was confirmed by comparison of the monoacetyldihydro- compounds, m.p. and mixed m.p. 146–147 °C. Reduction of the quinone by sulphur dioxide in methanol gave the quinol, identical with the dihydroxy compound (VIII) obtained by demethylation of the trimethoxyphenol with alcoholic hydrochloric acid(3), m.p. and mixed m.p. 242–243 °C.

(ii) *3-Hydroxy-1,2,4-Trimethoxy-10-Methylacridone* (VII).—The trimethoxyphenol from melicopidine when treated with nitric acid as described above, gave in good yield, the dimethoxy-

*o*-quinone (X), dark red needles from ethanol, m.p. 232–233 °C., mixed m.p. with a specimen prepared from melicopine 232–234 °C.\*

Found: C, 64.4; H, 4.2; N, 4.6; CH<sub>3</sub>O, 20.0%.

Calculated for C<sub>16</sub>H<sub>13</sub>O<sub>3</sub>N: C, 64.2; H, 4.3; N, 4.7; CH<sub>3</sub>O, 20.7%.

The identity was confirmed by reduction to the quinol, m.p. and mixed m.p. 163–164 °C. and by reductive acetylation to the monoacetyldihydro- compound, m.p. and mixed m.p. 156–158 °C. Moreover, the quinol was identical (m.p. and mixed m.p.) with the dihydroxy compound (XI) resulting from treatment of the trimethoxyphenol with alcoholic hydrochloric acid.

(iii) *1-Hydroxy-2-Ethoxy-3,4-Dimethoxy-10-Methylacridone* (XVII).—The ethoxydimethoxyphenol from melicopine was treated with nitric acid as above. The resulting *ethoxymethoxy-p-quinone* (XIX), obtained in good yield, crystallized from alcohol as bright red needles, m.p. 195–196 °C.

Found: C, 65.0; H, 4.7; N, 4.6; CH<sub>3</sub>O, 19.7%.

Calculated for C<sub>17</sub>H<sub>15</sub>O<sub>5</sub>N: C, 65.2; H, 4.8; N, 4.5; CH<sub>3</sub>O, 19.8% (one methoxyl and one ethoxyl calculated as methoxyl).

Reduction by sulphur dioxide in methanol gave the quinol, which melted at 200–201 °C. when alone or in admixture with the dihydroxy compound resulting from the action of alcoholic hydrochloric acid on the ethoxydimethoxyphenol(3). The latter was converted to the quinone (m.p. and mixed m.p. 195–196 °C.) by treatment with nitric acid. The quinone did not dissolve in cold aqueous sodium carbonate and was recovered unchanged (m.p. and mixed m.p.) after boiling an alcoholic solution for five minutes with *o*-phenylenediamine.

(iv) *3-Hydroxy-2-Ethoxy-1,4-Dimethoxy-10-Methylacridone* (XVIII).—The ethoxydimethoxyphenol from melicopine gave with nitric acid the *ethoxymethoxy-o-quinone* (XXI) in good yield. The substance crystallized from methanol as small, dark red needles, m.p. 225–226 °C.

Found: C, 65.2; H, 5.0; N, 4.5; CH<sub>3</sub>O, 19.4%.

Calculated for C<sub>17</sub>H<sub>15</sub>O<sub>6</sub>N: C, 65.2; H, 4.8; N, 4.5; CH<sub>3</sub>O, 19.8% (one methoxyl and one ethoxyl calculated as methoxyl).

The quinone did not dissolve in cold sodium carbonate solution. Reduction with sulphur dioxide in methanol gave the quinol, red needles, or orange-red prisms from aqueous methanol, m.p. 169–170 °C. The mixed m.p. with the dihydroxy- compound obtained(3) by the action of alcoholic acid on the ethoxydimethoxyphenol (XVIII) was 170–171 °C. For confirmation, the quinone was reductively acetylated in the usual manner and the product further acetylated with acetic anhydride and anhydrous sodium acetate to the diacetyldihydro- compound, m.p. and mixed m.p. with the diacetate of the dihydroxy compound 171–172 °C. The cycle was completed by the conversion of the dihydroxy compound (obtained from XVIII by alcoholic acid) to the quinone by treatment with nitric acid. The m.p. of the quinone was 224–225 °C.; a mixture with the specimen obtained directly from XVIII melted at 225–226 °C. The quinone reacted rapidly with *o*-phenylenediamine in alcoholic solution at room temperature giving a *phenazine*, yellow needles from ethyl acetate-light petroleum, m.p. 213–214 °C.

Found: C, 71.5; H, 5.0; N, 10.8%.

Calculated for C<sub>20</sub>H<sub>19</sub>O<sub>3</sub>N<sub>2</sub>: C, 71.7; H, 4.9; N, 10.9%.

#### (c) Action of Sodium Carbonate on the Ethoxymethoxyquinones

(i) *2-Ethoxy-3-Methoxy-10-Methylacridone-1,4-Quinone* (XIX).—The ethoxymethoxy-*p*-quinone, obtained from melicopine as described above, dissolved in hot aqueous sodium carbonate giving a purplish-red solution. Acidification afforded a mass of orange-red felted needles of the *hydroxymethoxyquinone* (XX). The substance separated from ethyl acetate as small red needles, m.p. 197–198 °C.

\* Preparation from the trimethoxyphenol is the most satisfactory method for the dimethoxy-*o*-quinone.



Found: C, 64.0; H, 4.4; N, 4.7;  $C_2H_5O$ , 15.7%.

Calculated for  $C_{16}H_{13}O_5N$ : C, 64.2; H, 4.3; N, 4.7;  $C_2H_5O$ , 15.1% (one ethoxyl).

Acetylation with acetic anhydride and pyridine gave a *monoacetate*, crimson needles from chloroform-ethyl acetate, m.p. 218–219 °C.

Found: C, 63.1; H, 4.5; N, 4.0;  $CH_3CO$ , 12.9%.

Calculated for  $C_{16}H_{12}O_4N(OOC.CH_3)$ : C, 63.3; H, 4.4; N, 4.1;  $CH_3CO$ , 12.6%.

Reduction of XX by sulphur dioxide in methanol, followed by dilution with water, yielded the *quinol*, dull red needles from aqueous methanol, m.p. 239–240 °C.

Found: C, 64.0; H, 5.0; N, 4.8%.

Calculated for  $C_{16}H_{15}O_5N$ : C, 63.8; H, 5.0; N, 4.7%.

The quinol dissolved in caustic soda to an intense violet solution. Reductive acetylation of the quinone or direct acetylation of the quinol with acetic anhydride and pyridine gave the *diacetyldihydro-* derivative, yellow needles from ethyl acetate-light petroleum, m.p. 188–189 °C.

Found: C, 62.4; H, 4.9; N, 3.7;  $CH_3CO$ , 22.3%.

Calculated for  $C_{16}H_{13}O_3N(OOC.CH_3)_2$ : C, 62.3; H, 4.9; N, 3.6;  $CH_3CO$ , 22.3%.

The quinone was refluxed with 46% hydrobromic acid for 30 minutes, the separated hydrobromide suspended in aqueous sodium bicarbonate and reduced by sodium hydrosulphite. The precipitated reduction product was acetylated with acetic anhydride and pyridine giving the triacetate of the dihydroxyquinol(2), yellow needles from ethyl acetate-light petroleum, m.p. and mixed m.p. 192–193 °C.

(ii) *2-Ethoxy-1-Methoxy-10-Methylacridone-3,4-Quinone* (XXI).—The ethoxymethoxy-*o*-quinone, obtained from melicopidine, likewise underwent demethylation on heating with aqueous sodium carbonate. The resulting hydroxyethoxyquinone was identical with XX obtained from the ethoxymethoxy-*p*-quinone, m.p. and mixed m.p. 197–198 °C.\* This was confirmed by comparison of the quinols (m.p. and mixed m.p. 239–240 °C.) and their diacetates (m.p. and mixed m.p. 188–189 °C.).

#### (d) Action of Nitrous Acid on *normelicopine* and *normelicopidine*

(i) *normelicopine* (5 g.) was triturated with 35% hydrochloric acid (100 ml.), cooled to 0 °C., and a solution of nitrous acid (1.06 mole: from 12.5 ml. 10% sodium nitrite) added gradually with stirring. When the *normelicopine* hydrochloride had reacted, the mixture was poured onto excess sodium bicarbonate, filtered, and made just acid with hydrochloric acid when a greenish-yellow crystalline solid separated. This substance, the hydroxymethoxyquinone (XIV), could not be purified by crystallization as it decomposed giving amorphous products when heated with neutral solvents or acetic acid. Moreover, it had no definite m.p. but darkened above 100 °C. partially liquefied and gradually decomposed above 130 °C. Consequently the substance could not be satisfactorily purified for analytical purposes. However, a sample dissolved in sodium bicarbonate, reprecipitated by acid, and washed with water was analysed after vacuum drying at room temperature.

Found: C, 58.8; H, 4.6%.

Calculated for  $C_{15}H_{11}O_5N.H_2O$ : C, 59.4; H, 4.3%.

The yield was c. 85% determined by precipitation of the quinol from bicarbonate solution with sodium hydrosulphite (*vide infra*).

Acetylation of the quinone (XIV) with acetic anhydride and pyridine gave a stable *monoacetate*, crimson needles from chloroform-ethyl acetate, m.p. 230–232 °C. (decomp.).

Found: C, 62.4; H, 4.2; N, 4.5;  $CH_3CO$ , 13.5%.

Calculated for  $C_{15}H_{10}O_4N(OOC.CH_3)$ : C, 62.4; H, 4.0; N, 4.3;  $CH_3CO$ , 13.1%.

\* This quinone is also identical (m.p. and mixed m.p. of the substance and its diacetyldihydro-derivative) with the acid,  $C_{16}H_{13}O_5N$ , obtained as a by-product from the action of alcoholic hydrochloric acid on the ethoxydimethoxyphenol (XVIII) from melicopidine, as described in Part II(3).

The quinone reacted readily with *o*-phenylenediamine in hot alcoholic solution giving a *phenazine*, violet-red needles from ethyl acetate, m.p. 281–283 °C. (uncorr.).

Found: C, 70.7; H, 4.3; N, 11.8;  $\text{CH}_3\text{O}$ , 9.3%.

Calculated for  $\text{C}_{21}\text{H}_{15}\text{O}_3\text{N}_3$ : C, 70.6; H, 4.2; N, 11.8;  $\text{CH}_3\text{O}$ , 8.7%.

The phenazine was recovered unchanged (m.p. and mixed m.p.) after refluxing with acetic anhydride and pyridine for one hour. Solutions of the quinone in non-hydroxylic solvents such as chloroform or benzene are olive-green. Its solutions in bicarbonate are deep brownish-red, becoming paler at high pH; in 5% caustic soda pale brownish-red, in 10% caustic soda almost yellow. The quinone is stable in acid solutions which are orange-red in colour, resembling solutions of the isomeric hydroxymethoxyquinone (XV). It is also stable in alkaline solution at room temperature and was recovered unchanged after passing oxygen for two hours through a solution in 5% caustic soda. It was identified by conversion to the diacetyldihydro- derivative (*vide infra*), m.p. 216–217 °C., mixed m.p. with an authentic specimen 217–218 °C.

The quinone (XIV) was reduced by sulphur dioxide in methanol to the *quinol*, orange-red needles, m.p. 176–178 °C. from aqueous methanol containing a little sulphur dioxide.

Found: C, 62.2; H, 4.5; N, 4.9;  $\text{CH}_3\text{O}$ , 11.1%.

Calculated for  $\text{C}_{16}\text{H}_{13}\text{O}_5\text{N}$ : C, 62.7; H, 4.5; N, 4.9;  $\text{CH}_3\text{O}$ , 10.8% (one methoxyl).

Reduction to the quinol was also effected by addition of sodium hydrosulphite to a bicarbonate solution of the quinone. The quinol was reoxidized to the quinone by passing oxygen through a solution in 2% caustic soda for one hour. Acidification gave a greenish-yellow precipitate which was acetylated to the quinone monoacetate, m.p. and mixed m.p. 230–232 °C.

Acetylation of the quinol gave the *diacetate*, orange prisms from methanol, m.p. 218–219 °C. The same diacetate (m.p. and mixed m.p.) resulted from reductive acetylation of the quinone (XIV).

Found: C, 61.6; H, 4.6; N, 4.0;  $\text{CH}_3\text{CO}$ , 23.5%.

Calculated for  $\text{C}_{18}\text{H}_{11}\text{O}_3\text{N}(\text{OOC}\cdot\text{CH}_3)_2$ : C, 61.5; H, 4.6; N, 3.8;  $\text{CH}_3\text{CO}$ , 23.2%.

Further acetylation with acetic anhydride and anhydrous sodium acetate gave the *triacetate*, yellow needles from ethyl acetate-light petroleum, m.p. 208.5–209.5 °C.

Found: C, 61.3; H, 4.8; N, 3.5;  $\text{CH}_3\text{CO}$ , 31.3%.

Calculated for  $\text{C}_{18}\text{H}_{10}\text{O}_5\text{N}(\text{OOC}\cdot\text{CH}_3)_3$ : C, 61.0; H, 4.6; N, 3.4;  $\text{CH}_3\text{CO}$ , 31.2%.

Reductive methylation of the quinone with dimethyl sulphate and caustic soda containing sodium hydrosulphite gave melicopicine in 95% yield, m.p. and mixed m.p. with an authentic specimen 132–133 °C.

Found: C, 65.8; H, 5.9; N, 4.3%.

Calculated for  $\text{C}_{18}\text{H}_{10}\text{O}_5\text{N}$ : C, 65.7; H, 5.8; N, 4.3%.

The identity of the base was further confirmed by conversion to *normelicopicine*, m.p. and mixed m.p. 128–129 °C.

Found: C, 64.9; H, 5.6; N, 4.5;  $\text{CH}_3\text{O}$ , 29.8%.

Calculated for  $\text{C}_{17}\text{H}_{17}\text{O}_5\text{N}$ : C, 64.8; H, 5.4; N, 4.4;  $\text{CH}_3\text{O}$ , 29.5%.

The quinone was refluxed with 46% hydrobromic acid for 30 minutes, the product suspended in bicarbonate solution and reduced with sodium hydrosulphite. Acetylation of the insoluble dihydro- compound gave the triacetate of the dihydroxyquinol, m.p. and mixed m.p. 192–194 °C.

Although stable in alkaline solution at room temperature, the quinone (XIV) is decomposed by hot alkali. After heating on the water bath for  $\frac{1}{2}$ –1 hour with 10% caustic soda, the mixture (from which a solid had separated) was cooled and extracted with chloroform. Evaporation of the chloroform gave *1-methyl-3-methoxyacetyl-4-quinolone*, shining colourless needles from methanol or water, m.p. 208–210 °C. after sintering at 205 °C.

Found: C, 67.1; H, 5.8; N, 6.0; O, 20.6;  $\text{CH}_3\text{O}$ , 13.4%.

Calculated for  $\text{C}_{13}\text{H}_{13}\text{O}_5\text{N}$ : C, 67.5; H, 5.6; N, 6.1; O, 20.8;  $\text{CH}_3\text{O}$ , 13.4%.

The substance reacted with 2,4-dinitrophenylhydrazine to give an orange-red microcrystalline 2,4-dinitrophenylhydrazone which was not analysed.

Oxidation of the base with acid permanganate gave I, identified by m.p. and mixed m.p. 294–296 °C. (uncorr.).

(ii) *nor*Melicopidine was treated with nitrous acid under the conditions employed for *nor*-melicopine. If the strongly acid reaction mixture was allowed to stand for one–two hours, a dark red crystalline hydrochloride separated, which consisted essentially of the hydrochloride of the dihydroxyquinone (IV), identified by the m.p. and mixed m.p. of the dihydro- and triacetyl-dihydro- derivatives. Consequently the mixture was not allowed to stand for more than five minutes after addition of the nitrous acid before pouring onto bicarbonate. The green bicarbonate solution on acidification afforded an orange-yellow precipitate of the 1-methoxy-2-hydroxy-3,4-quinone (XVI). This precipitate redissolved to a green solution if basified immediately with bicarbonate, but was unstable in contact with acids, being converted to a phenol not identical with the quinol. It was therefore necessary to filter the precipitated quinone rapidly and to use at once without drying. The yield was low. However, treatment of the bicarbonate solution with sodium hydrosulphite precipitated the quinol in 45–50% yield. The quinone decomposed on attempted crystallization from a variety of solvents and was not analysed. It reacted with acetic anhydride and pyridine, and with *o*-phenylenediamine, but the product in each case was a mixture which could not be purified by crystallization. The *quinol*, which was prepared by reduction of the quinone in bicarbonate solution with hydrosulphite, and also by sulphur dioxide in methanol, separated from aqueous methanol as orange-red needles, m.p. 227–228 °C.

Found: C, 62.5; H, 4.4; N, 4.9; CH<sub>3</sub>O, 11.2%.

Calculated for C<sub>15</sub>H<sub>13</sub>O<sub>5</sub>N: C, 62.7; H, 4.5; N, 4.9; CH<sub>3</sub>O, 10.8% (one methoxyl).

Reductive acetylation of the quinone or direct acetylation of the quinol gave the *diacetyl-dihydro-* compound, yellow needles from ethyl acetate–light petroleum, m.p. 193–194 °C.\*

Found: C, 61.4; H, 4.6; N, 3.8; CH<sub>3</sub>O, 8.4; CH<sub>3</sub>CO, 23.2%.

Calculated for C<sub>15</sub>H<sub>11</sub>O<sub>3</sub>N(OOC.CH<sub>3</sub>)<sub>2</sub>: C, 61.5; H, 4.6; N, 3.8; CH<sub>3</sub>O, 8.4; CH<sub>3</sub>CO, 23.2%.

Further acetylation with acetic anhydride and anhydrous sodium acetate gave the *triacetate*, yellow prisms from ethanol, m.p. 167.5–168.5 °C.

Found: C, 61.3; H, 4.9; N, 3.5; CH<sub>3</sub>CO, 30.7%.

Calculated for C<sub>15</sub>H<sub>10</sub>O<sub>2</sub>N(OOC.CH<sub>3</sub>)<sub>3</sub>: C, 61.0; H, 4.6; N, 3.4; CH<sub>3</sub>CO, 31.2%.

Reductive methylation of XVI gave melicopine, m.p. and mixed m.p. 132–133 °C., and treatment with 46% hydrobromic acid gave the dihydroxyquinone (IV), identified as the triacetate of the quinol, m.p. and mixed m.p. 191–193 °C.

#### (e) Action of Caustic Soda on the Dialkoxyquinones

(i) The dimethoxy-*p*-quinone (IX) when treated with cold 5% caustic soda dissolved at once to a brownish-red solution which was immediately acidified with 10% hydrochloric acid. A greenish-yellow solid separated, resembling the hydroxyquinone (XIV). It was identified as such by means of its acetyl- and diacetyldihydro- derivatives, the former melting at 229–231 °C. and the latter at 217–218 °C. In neither case was the melting point depressed by admixture with an authentic specimen.

Found for the diacetyldihydro- derivative: C, 61.4; H, 4.7; CH<sub>3</sub>CO, 23.5%.

Calculated for C<sub>15</sub>H<sub>11</sub>O<sub>3</sub>N(OOC.CH<sub>3</sub>)<sub>2</sub>: C, 61.5; H, 4.6; CH<sub>3</sub>CO, 23.2%.

XIV was also obtained (m.p. and mixed m.p. of the acetate) after a rapid stream of oxygen had been passed for two hours through a solution of the dimethoxy-*p*-quinone in caustic soda.

The ethoxymethoxy-*p*-quinone (XIX) behaved in the same manner when treated with cold caustic soda, giving the hydroxyquinone (XIV), identified as the diacetyldihydro- derivative, m.p. and mixed m.p. 216–217 °C.

\* A mixed m.p. with the triacetyldihydro- derivative of the dihydroxyquinone (m.p. 193–194 °C.) was depressed 10–15 °C.



(ii) Both the dimethoxy-*o*-quinone (X) and the ethoxymethoxy-*o*-quinone (XXI) dissolved in cold 5% caustic soda to give yellow solutions (agitation was necessary to prevent local concentration gradients which affected the course of the reaction). On addition of bicarbonate the colour of both solutions changed to green (compare the colour of solutions of XIV at different pH values). Acidification afforded some orange-yellow unstable solid. The amount separating was low, so identification was effected by reduction of the bicarbonate solution of the quinone with sodium hydrosulphite. The resulting orange-red precipitate (yield c. 15–20%) was identified in both cases as the quinol corresponding to XVI by preparation of the diacetyldihydro-derivatives, m.p. and mixed m.p. 192–194 °C.

Oxidation took place when a stream of oxygen was passed through a solution of the dialkoxymethoxy-*o*-quinones in caustic soda. The product was largely phenolic and was not further investigated; neither the quinone (XVI) nor the acid,  $C_{18}H_8O_6N$  (XLI), could be detected.

(f) *Action of Ethanolic Potash on the Alkaloids*

(i) *Melicopicine*.—The alkaloid (10 g.) in ethanol (750 ml.) containing potassium hydroxide (45 g.) was refluxed for 24 hours. The bulk of the solvent was then evaporated, water (400 ml.) added, and the remainder of the ethanol removed.

The mixture was extracted repeatedly with chloroform and the combined chloroform extracts evaporated leaving an oily residue (c. 4 g.). The alkaline solution was saturated with carbon dioxide and extracted with chloroform which on evaporation left c. 3 g. yellow phenolic material. Bicarbonate-soluble by-products were not examined. The phenol (XXX) crystallized from ethyl acetate-light petroleum as pale yellow needles, m.p. 176.5–177.5 °C.

Found: C, 64.7; H, 5.4; N, 4.3;  $CH_3O$ , 29.4%.

Calculated for  $C_{17}H_{17}O_8N$ : C, 64.8; H, 5.4; N, 4.4;  $CH_3O$ , 29.5%.

Acetylation with acetic anhydride and pyridine gave the *acetate*, pale yellow needles from chloroform-light petroleum, m.p. 130–131 °C.

Found: C, 63.7; H, 5.4; N, 4.0;  $CH_3CO$ , 12.0%.

Calculated for  $C_{17}H_{16}O_4N(OOC.CH_3)$ : C, 63.9; H, 5.3; N, 3.9;  $CH_3CO$ , 12.0%.

Methylation of the phenol with dimethyl sulphate and caustic soda gave melicoprene, m.p. and mixed m.p. 133–134 °C. The phenol was recovered unchanged (m.p. and mixed m.p.) after refluxing for two hours with 10% alcoholic hydrochloric acid.

The alkali-insoluble fraction was dissolved in benzene and the solution extracted three times with 20% hydrochloric acid. The benzene was washed with water and evaporated to dryness. The residue (XXXII; 1.15 g.) crystallized from aqueous methanol as orange-red needles, m.p. 88–89 °C.

Found: C, 65.7; H, 5.8; N, 4.1;  $CH_3O$ , 27.9%.

Calculated for  $C_{18}H_{18}O_5N$ : C, 65.7; H, 5.8; N, 4.3;  $CH_3O$ , 28.3% (two methoxyls and one ethoxyl calculated as methoxyl).

This substance was treated with 68% nitric acid, the mixture poured onto excess sodium bicarbonate and extracted with chloroform. The residue after evaporation of the chloroform crystallized from ethanol as bright red needles, m.p. 195–196 °C. alone or in admixture with the ethoxymethoxy-*p*-quinone (XIX).

The 20% hydrochloric acid extracts deposited, on standing, a mass of fine orange needles which were filtered, treated with alkali, and the base shaken into chloroform. Evaporation of the chloroform left a yellow gum which would not crystallize. Consequently it was converted to the *nore* compound by treatment with alcoholic hydrochloric acid. The product was identical (m.p. and mixed m.p. 87–88 °C.) with the substance (XXXII) which was not removed from the benzene by 20% hydrochloric acid (*vide supra*).\*

\* The non-crystallizable base was presumably 2-ethoxy-1,3,4-trimethoxy-10-methylacridone (XXXI) which should also result from the methylation of, for example, the 1-hydroxy-2-ethoxy-3,4-dimethoxy- compound (XVII) or the 2-ethoxy-1,3,4-trihydroxy- compound. Both these substances were methylated and in each case the product failed to crystallize.



(ii) *Melicopine*.—The combined alkali-insoluble material from several preparations of the ethoxydimethoxyphenol (XVII) from melicopine was separated by hot ethanol into two fractions. The less soluble fraction was identified as *normelicopine* (m.p. and mixed m.p.) after recrystallization from chloroform-ethanol.

The more soluble fraction was dissolved in benzene and submitted to repeated fractional extraction with 2½, 5, and 10% hydrochloric acid. The 2½ and 5% acid extracts contained mainly melicopine (m.p. and mixed m.p.). The base from the 10% acid extracts was twice precipitated by ether from chloroform giving some more melicopine (ether-insoluble) together with a substance which could not be further purified and which crystallized from ethanol as yellow needles, m.p. 151–153 °C., mixed m.p. with melicopine 152–174 °C. On refluxing with 5% ethanolic hydrochloric acid, this substance (225 mg.) gave *normelicopine* (200 mg.) m.p. and mixed m.p. 234–236 °C.

#### (g) Reduction of Melicopine

When melicopine was heated on the water-bath with a large excess of zinc dust and concentrated hydrochloric acid, the colour of the solution gradually became paler and a bright yellow solid separated. While the reduction was in progress the mixture was markedly susceptible to aerial oxidation as shown by the development of a deep red colour at the exposed surfaces. After cooling and dilution with water, the mixture was shaken with chloroform giving a greenish-yellow extract. Evaporation of the chloroform left a dark coloured residue which was treated with boiling alcohol leaving an insoluble yellow solid. Crystallization from benzene-light petroleum afforded small yellow fluorescent prisms, m.p. 333–334 °C. (uncorr.).

Found: C, 68.8; H, 5.1; N, 4.6; CH<sub>3</sub>O, 21.4%.

Calculated for C<sub>34</sub>H<sub>30</sub>O<sub>8</sub>N<sub>2</sub>: C, 68.7; H, 5.1; N, 4.7; CH<sub>3</sub>O, 20.9% (four methoxyls).

The substance was almost insoluble in alcohol, insoluble in alkalis, and only sparingly soluble in strong hydrochloric acid. It dissolved in boiling dilute nitric acid, undergoing oxidation, and giving an intense reddish-violet solution from which crystals of the diacridinium dinitrate separated on cooling. These were dissolved in water and treated with saturated aqueous picric acid. The diacridinium dipicrate, recrystallized from methanol, was obtained as glistening violet plates, m.p. 206–207 °C.

Found: C, 52.8; H, 3.6; N, 10.7%.

Calculated for C<sub>34</sub>H<sub>30</sub>O<sub>8</sub>N<sub>2</sub><sup>++</sup>.2(C<sub>6</sub>H<sub>2</sub>O<sub>7</sub>N<sub>3</sub>)<sup>-</sup>: C, 52.6; H, 3.2; N, 10.7%.

In common with other diacridinium compounds, 2,2', 3,3', and 4,4'-dimethoxydiacridinium salts exhibit chemiluminescence when treated with alkaline hydrogen peroxide, but Gleu and Nitzsche(17) found that the 1,1'-dimethoxy- compound did not. The diacridinium dinitrate from melicopine did not exhibit chemiluminescence with ammoniacal hydrogen peroxide.

Test-tube experiments showed that melicopidine and melicopine were also reduced by zinc and hydrochloric acid. The reaction, however, was complex and the isolation of the diacridenes was not attempted.

#### (h) Preparation and Properties of the Acid, C<sub>13</sub>H<sub>9</sub>O<sub>6</sub>N (XLI)

The preparation of this acid from the 1,2-dialkoxy-3,4-dihydroxy-10-methylacridones has been described in Part II(3). A more convenient method is as follows. Melicopidine dissolved in concentrated hydrochloric acid (20 ml./g. alkaloid) was treated at 0 °C. with nitrous acid (10 moles). After standing overnight the white solid, which had separated, was filtered and dissolved in cold aqueous sodium bicarbonate. The filtered solution was acidified giving a white crystalline precipitate, m.p. and mixed m.p. with the acid, C<sub>13</sub>H<sub>9</sub>O<sub>6</sub>N, 236–238 °C. (decomp.). The yield was c. 50%.

Distillation of the acid, well mixed with ten times its weight of zinc dust, yielded a colourless crystalline base. This was converted to its picrate, a yellow microcrystalline powder from methanol, m.p. 229–231 °C., mixed m.p. with the picrate of 1-methyl-4-quinolone 230–231 °C. To a solution of the acid (1 g.) in a mixture of water (10 ml.) and sulphuric acid (98%; 15ml.), saturated aqueous potassium permanganate was added until the colour persisted. The solution

was cooled, rendered nearly neutral by the cautious addition of concentrated aqueous ammonia, and the precipitate (0.46 g.) filtered. Crystallization from acetic acid gave colourless needles which shrank at 280 °C. and melted at 294–295 °C. (uncorr.) alone or mixed with 1-methyl-4-quinolone-3-carboxylic acid (I). Oxidation with 68% nitric acid as described in Part IV(1) for the alkaloids also gave I in good yield (m.p. and mixed m.p.). The acid (XLI) was decarboxylated by heating at 180–200 °C. for 10–15 minutes in dibutyl phthalate containing a little copper-bronze. The pale brown solution was cooled, diluted with chloroform (3 vol.), and extracted once with aqueous sodium carbonate and then three times with 20% hydrochloric acid. Acidification of the sodium carbonate extract precipitated a small amount of a colourless acid. Recrystallization from acetic acid gave I, m.p. and mixed m.p. 294–296 °C. (uncorr.) after shrinking at 280 °C. The combined acid extracts were basified with ammonia and repeatedly extracted with chloroform. The residue after evaporation of the chloroform crystallized from chloroform-light petroleum as colourless needles, m.p. 217–218 °C.

Found: C, 70.6; H, 4.8; N, 7.6%.

Calculated for  $C_{11}H_9O_2N$ : C, 70.6; H, 4.8; N, 7.5%.

The base is appreciably soluble in water and readily soluble in dilute acids. It was oxidized by cold dilute acid permanganate to I (m.p. and mixed m.p.). With Brady's reagent it formed a 2,4-dinitro-phenylhydrazone, small crimson needles from acetic acid, m.p. 318–319 °C. (uncorr.)

Found: C, 55.7; H, 3.6; O, 22.2%.

Calculated for  $C_{17}H_{13}O_5N_5$ : C, 55.6; H, 3.5; O, 21.8%.

The acid (XLI) was recovered unchanged after several hours' refluxing with either 10% caustic soda or 20% hydrochloric acid.

#### (i) Preparation and Properties of the Acid, $C_{13}H_9O_5N$ (XLII)

2-Hydroxy-3-methoxy-10-methylacridone-1,4-quinone (XIV, from 5 g. *normeleopine*) dissolved in caustic soda, was oxidized at 0 °C. with potassium permanganate. Excess permanganate and the precipitated manganese dioxide were removed by addition of sodium sulphite and acidification. On boiling the acid solution a white crystalline solid (3.33 g.) separated. Crystallization from acetic acid yielded colourless needles, m.p. 254–256 °C. (uncorr.) after shrinking at 245 °C.\*

Found: C, 60.4; H, 3.5; N, 5.2%.

Calculated for  $C_{13}H_9O_5N$ : C, 60.2; H, 3.5; N, 5.4%.

In the preparation it is not necessary to isolate the hydroxymethoxyquinone: the oxidation can be carried out on the reaction mixture after treatment of *normeleopine* with nitrous acid. The acid,  $C_{13}H_9O_5N$ , was identical (m.p. and mixed m.p.) with a by-product from the aerial oxidation of the dimethoxy *o*-quinol (XI) described in Part II(3). XLII was also prepared by reduction of the acid,  $C_{13}H_9O_6N$  (XLI), which was dissolved in 10% caustic soda and heated on the water-bath for five hours with zinc dust. Acidification of the filtered solution and crystallization of the precipitate from acetic acid afforded colourless needles, m.p. and mixed m.p. 254–256 °C. (uncorr.) after shrinking at 245 °C. Oxidation of XLI to XLII was attempted but without success. Oxidation of XLII with boiling 68% nitric acid gave 1-methyl-4-quinolone-3-carboxylic acid, identified by m.p. and mixed m.p.

Decarboxylation was carried out by heating at 200–220 °C. for 10 minutes in dibutyl phthalate (containing a little copper bronze). The solution became intense reddish-violet, which changed to red, and developed a strong yellow fluorescence. It was cooled, diluted with chloroform (4 vol.), and extracted once with 10% hydrochloric acid, and three times with concentrated hydrochloric acid. The 10% acid extract contained only a little impure material and was discarded. The strong acid solutions were basified with ammonia (with cooling) and the precipitate shaken into chloroform. Evaporation of the chloroform left a brown crystalline residue, which after several

\* In all m.p. determinations this substance turned deep violet between 180 and 200 °C. Above 200 °C. the colour changed to reddish-brown.

crystallizations from chloroform-alcohol afforded flat yellow needles, m.p. 273–274 °C. (uncorr.). The substance was obtained colourless by short treatment with hot alcoholic hydrochloric acid (5%). The m.p. of the product, colourless needles from chloroform-alcohol, was unchanged and there was no depression of the m.p. of a mixture with the yellow material.

Found: C, 67.1; H, 4.1; N, 6.6%.

Calculated for  $C_{12}H_9O_3N$ : C, 67.0; H, 4.2; N, 6.5%.

The substance is a very weak base. It dissolved slowly in hot 20% caustic soda and a white solid separated. This solid, presumably a sodium salt, was easily soluble in cold water and no immediate precipitation took place on acidification of the aqueous solution. On standing, the original base (m.p. and mixed m.p.) separated.

#### (j) Preparation and Properties of the Acid, $C_{13}H_9O_4N$ (XL)

The most convenient method of preparation is as follows: *nor*Melicopine triturated with concentrated hydrochloric acid (20 ml./g.) was treated at 0 °C. with nitrous acid (5 moles). After standing for an hour, a pale yellow gelatinous mass had separated from the reaction mixture. Excess nitrous acid was destroyed by the addition of urea, the mixture basified with caustic soda and shaken in air until the green colour disappeared. Acidification and boiling precipitated XL, m.p. and mixed m.p. 261–263 °C. (uncorr.) after crystallization from acetic acid. Yield, 45%. The yellow gelatinous intermediate is considered to be identical with the substance obtained by the action of nitric acid on the dihydroxyquinone(2). XL is a weak acid, being slowly extracted by chloroform from a solution in bicarbonate.

It was recovered unchanged (m.p. and mixed m.p.) after refluxing with acetic anhydride and pyridine, and after refluxing with alcoholic 2,4-dinitrophenylhydrazine hydrochloride. The results of attempted decarboxylation at 260–270 °C. in dibutyl phthalate suggested the absence of a carboxyl group. The acid was not recovered, but underwent decomposition and a small amount of intractable gummy basic material resulted. The only identifiable product was a small amount of impure 1-methyl-4-quinolone-3-carboxylic acid (I).

Oxidation of XL with either boiling 68% nitric acid or hot alkaline hydrogen peroxide gave I, m.p. and mixed m.p. 294–296 °C. (uncorr.) after shrinking at 280 °C. Oxidation with alkaline permanganate at 0 °C. gave the acid  $C_{13}H_9O_6N$  (XLI), m.p. and mixed m.p. 236–238 °C. (decomp.).

Found: C, 57.1; H, 3.4; N, 5.3%.

Calculated for  $C_{13}H_9O_6N$ : C, 56.7; H, 3.3; N, 5.1%.

### III. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Industrial Chemistry, C.S.I.R.O. The authors are indebted to A. G. Pulford and A. Walsh, who examined and reported on the infra-red absorption spectra referred to in this paper, to Dr. H. H. Hatt for valuable criticism of the manuscript of this and the preceding papers of the series, to R. G. Cooke for helpful discussion of several points which arose during the course of the work, and to Professor E. J. Hartung for providing accommodation and facilities in the Chemistry Department, University of Melbourne.

### IV. REFERENCES

- (1) PRICE, J. R.—*Aust. J. Sci. Res. A* **2**: 272 (1949).
- (2) CROW, W. D.—*Ibid.* **2**: 264 (1949).
- (3) CROW, W. D., and PRICE, J. R.—*Ibid.* **2**: 255 (1949).
- (4) STOCKELBACH, F. E.—U.S. Pat. 1,792,716 and 1,792,717 (February 17, 1931). (*Chem. Abstr.* **25**: 2154 (1931).)



- (5) ONO, K., and IMOTO, M.—*J. Chem. Soc. Japan* **59** : 359 (1938) ; and earlier papers.
- (6) ROBINSON, G. M., and ROBINSON, R.—*J. Chem. Soc.* **111** : 929 (1917).
- (7) PARIJS, A. H.—*Rec. Trav. chim. Pays-Bas* **49** : 33 (1930).
- (8) BERINZAGHI, B., DEULOFEU, V., LABRIOLA, R., and MURUZABAL, A.—*J. Amer. Chem. Soc.* **65** : 1357 (1943).
- (9) HODGSON, H. H., and HABESHAW, J.—*J. Chem. Soc.* **1942** : 45 (1942).
- (10) FOX, H. H., and BOGERT, M. T.—*J. Amer. Chem. Soc.* **63** : 2996 (1941).
- (11) SLOOFF, G.—*Rec. Trav. chim. Pays-Bas* **54** : 995 (1935).
- (12) BÖESEKEN, J., and SLOOFF, G.—*Proc. Acad. Sci. Amst.* **37** : 584 (1934). (*Chem. Abstr.* **29** : 2937 (1935).)
- (13) SALWAY, A. H.—*J. Chem. Soc.* **97** : 2413 (1910).
- (14) SCHWENK, E., and PAPA, D.—*J. Org. Chem.* **10** : 232 (1945).
- (15) BIRCH, A. J.—*J. Chem. Soc.* **1947** : 102 (1947).
- (16) ALBERT, A. A., and GOLDACRE, R.—*Ibid.* **1943** : 452 (1943).
- (17) GLEU, K., and NITZSCHE, S.—*J. prakt. Chem.* **153** : 233 (1939).
- (18) KAUFFMANN, H., and FRITZ, I.—*Ber. dtsh. chem. Ges.* **43** : 1215 (1910).
- (19) FLETT, M. St. C.—*J. Chem. Soc.* **1948** : 1441 (1948).
- (20) SCHERKS, E.—*Ber. dtsh. chem. Ges.* **18** : 378 (1885).
- (21) GRAEBE, C., and TRÜMPY, F.—*Ibid.* **31** : 373 (1898).
- (22) NIETZKI, R., and BENCKISER, T.—*Ibid.* **19** : 293 (1886).



# GRAPHICAL STUDY OF THE DISPERSION OF ELECTRO-MAGNETO-IONIC WAVES

By V. A. BAILEY\* and J. A. ROBERTS†

[Manuscript received June 6, 1949]

## Summary

A graphical method for approximating to all the eight roots of the equation of dispersion, corresponding to any numerically specified case, is described.

This method uses curves drawn with  $\omega$  and  $l$  as coordinates to give readily the following information about the waves which can exist in the medium : (i) the frequency-bands in which undamped waves or wave-groups can grow as they progress, (ii) the wave-number bands in which unattenuated waves can grow in time, (iii) the phase- and group-velocities, refractive indices, and coefficients of positive and negative attenuation and damping, (iv) the general effect of collisions between electrons and other particles on the attenuation or damping of a wave or wave-group.

Several illustrative examples are given. The same method is also applied to a special case of the more comprehensive equation of dispersion which includes the effects due to the motions of the positive ions, and it is shown that there can then exist unattenuated waves which grow with the lapse of time.

## I. INTRODUCTION

In the theory of plane waves of the form  $\exp i(\omega t - lx)$  in an ionized medium, pervaded by static electric and magnetic fields, which has been given by one of us(1, 2, 3), the equation of dispersion (when the positive ions are assumed motionless) is of the eighth degree both in the angular frequency  $\omega$  and in the angular wave-number  $l$ . It is therefore *in general* necessary to determine the roots of an irreducible equation of the eighth degree, and even in some of the special cases of importance equations of the fourth degree are unavoidable.

We here describe a graphical method for approximating to all the roots of this equation which requires the solution of at the most a cubic equation. This method has been found to be both practicable and convenient when the medium is numerically specified.

This method is also applied to an important special case of the more comprehensive theory in which the motions of the positive ions are also taken into account.

The equation of dispersion can be conveniently expressed in the following form :

$$X(Y^2 - \Omega_i^2 Z^2) + U_T^2 l^2 [Y(R - iv) - \Omega_i^2 Z] - RY\Omega_T^2 Z - 2l^2 R\Omega_i Z(\check{\Omega}_T \cdot U_T) = 0, \dots (1)$$

\* Department of Physics, University of Sydney.

† Commonwealth Research Student, University of Sydney.

where

$$\left. \begin{aligned} X &= R^2 - \tau l^2 - 1 - i\nu R, \\ Y &= RZ + R - i\nu Z, \\ Z &= l^2 - \omega^2, \\ R &= \omega - U_1 l, \end{aligned} \right\} \dots\dots\dots (2)$$

$U_0$  = the mean drift velocity of the electrons,

$H_0$  = the static magnetic force,

$\tilde{\Omega}_0 = -H_0 e/mc$  = the gyro-frequency vector,

$\nu$  = frequency of collision of an electron with other kinds of particles,

$\tau$  = one-third of the mean square velocity of agitation of the electrons,

$N_0$  = the mean density of the electrons,

$e, m$  = the charge and mass of an electron, respectively;

the subscripts 1 and  $T$  respectively denote components along and transverse to  $Ox$  the direction of propagation, and the units of velocity and angular frequency are taken as equal respectively to  $c$ , the velocity of light, and  $p$ , the electron density frequency, which is given by  $\sqrt{4\pi N_0 e^2 m}$ .

Since (1) is non-relativistic we must have  $U_T^2 \ll 1$ . It can be verified that the terms in  $U_T^2$  are then all negligible compared with certain terms in  $X(Y^2 - \Omega_1^2 Z^2)$ . We will therefore omit the terms in  $U_T^2$  from (1) in all that follows.

Three *physical* aspects of the dispersion equation will be considered, namely.

*Aspect 1.*—The possible waves with a given real phase velocity  $\omega/l$ , i.e. with a real refractive index  $M$  (A1 waves).

*Aspect 2.*—The possible waves with a given real frequency  $\omega$  (A2 waves).

*Aspect 3.*—The possible waves with a given real wave number  $l$  (A3 waves).

## II. THE APPROXIMATE DISPERSION EQUATION FOR SMALL COLLISION FREQUENCIES

When the terms in  $\nu^2$  may be neglected, (1) can be written as

$$P(\omega, l) + i\nu Q(\omega, l) = 0, \dots\dots\dots (3)$$

where

$$P = X_0[Y_0^2 - \Omega_1^2 Z^2] - R_0 Y \Omega_T^2 Z - 2l R \Omega_1 Z (\tilde{\Omega}_T \cdot \mathbf{U}_T) \dots\dots\dots (4a)$$

$$Q = -R[Y_0^2 - \Omega_0^2 Z^2 + 2X_0 Z(Z+1)] \dots\dots\dots (4b)$$

$$X_0 = R^2 - \tau l^2 - 1, \quad Y_0 = R(Z+1) \dots\dots\dots (5)$$

Since  $\nu$  is small, the solutions of

$$P(\omega, l) = 0 \dots\dots\dots (6)$$

may be taken as first approximations, and improved values found by Newtonian iteration.\*  $P$  is of the eighth degree in  $\omega$  and  $l$  and, in general, is irreducible.

For A2 waves, if  $l_0$  denotes a solution of (6), the second approximation is

$$l = l_0 - i\eta, \dots\dots\dots (7)$$

where

$$\eta = \nu Q(\omega, l_0) [\partial P(\omega, l_0) / \partial l_0]^{-1} \dots\dots\dots (8)$$

\* This process can, if necessary, also be applied to the more exact equation (1).

For 13 waves, denoting the first approximations by  $\omega_0$ , the second approximations are

$$\omega = \omega_0 - i\zeta, \quad \dots \dots \dots (9)$$

where

$$\zeta = \nu Q(\omega_0, l) [\partial P(\omega_0, l) / \partial \omega_0]^{-1} \dots \dots \dots (10)$$

### III. THE CURVE $P=0$

$P$  may conveniently be transformed by using the variables  $\omega$  and  $\sigma = l/\omega$ . With this notation

$$P(\omega, l) = \omega^2 P_1(\omega^2, \sigma), \quad \dots \dots \dots (11)$$

where

$$P_1(\omega^2, \sigma) = \omega^6 \varphi_6(\sigma) + \omega^4 \varphi_4(\sigma) + \omega^2 \varphi_2(\sigma) + \varphi_0, \quad \dots \dots \dots (12)$$

$$\left. \begin{aligned} \varphi_6 &= r^2 \zeta^2 (r^2 - \tau \sigma^2), \\ \varphi_4 &= r^2 \zeta (2r^2 - \Omega_0^2 \zeta - \zeta) - \tau \sigma^2 \zeta (2r^2 - \Omega_1^2 \zeta), \\ \varphi_2 &= r^2 (r^2 - 2\zeta - \Omega_2^2 \zeta - \tau \sigma^2) + \Omega_1^2 \zeta^2 - 2\sigma r \zeta (\Omega_1 (\tilde{\Omega}_1 \cdot \mathbf{U}_T)), \\ \varphi_0 &= -r^2. \end{aligned} \right\} \dots \dots \dots (13)$$

and

$$r = R/\omega = 1 - U_1 \sigma; \quad \zeta = Z/\omega^2 = \sigma^2 - 1 \quad \dots \dots \dots (14)$$

For any selected value of  $\sigma$ , the corresponding values of  $\omega$  satisfying  $P=0$  may be obtained by solving an equation (in  $\omega^2$ ) which is not of higher degree than three. In this way the curve  $P(\omega, l)=0$  can be plotted relatively easily. The solution of the cubic may conveniently be determined by means of tables (e.g. Jahneke and Emde(4)). Since only real values of  $\omega$  are required for this purpose, they may also be obtained by plotting the curve  $(x, y)$  where  $x = \omega^2$  and  $y = P_1(x, \sigma)$ .

#### (a) The Asymptotes and Intercepts on the Axes

The equations of the asymptotes are(5)

$$l = \sigma_n \omega + \delta_n \quad (n = \pm 1, \pm 2, \pm 3, \pm 4).$$

where  $\sigma_n$  are the roots of  $\varphi_6=0$ , namely,

$$\left. \begin{aligned} \sigma_{\pm 1} &= U_1^{-1}, \\ \sigma_{\pm 2} &= 1, \\ \sigma_{\pm 3} &= -1, \\ \sigma_{\pm 4} &= (U_1 \pm \sqrt{\tau})^{-1}, \end{aligned} \right\} \dots \dots \dots (15)$$

and  $\delta_n$  are the roots of(6)

$$\left. \begin{aligned} \frac{1}{2} \varphi_6'' \delta_n^2 + \varphi_5' \delta_n + \varphi_4 &= 0 \\ n &= \pm 1, \pm 2, \pm 3, \end{aligned} \right\} \dots \dots \dots (16)$$

for

and of (5)

$$\left. \begin{aligned} \varphi_6' \delta_n + \varphi_5 &= 0 \\ n &= \pm 4. \end{aligned} \right\} \dots \dots \dots (17)$$

for

$\varphi_5$  is the coefficient of  $\omega^5$  (which here is zero) and the prime denotes differentiation with respect to  $\sigma$ .

Thus

$$\left. \begin{aligned} \delta_{\pm 1} &= \pm \Omega_1 U_1^{-1}, & (\text{if } \tau \neq 0); \\ \delta_{\pm 2} &= 0, & (\text{if } \tau \neq (1 - U_1)^2); \\ \delta_{\pm 3} &= 0, & (\text{if } \tau \neq (1 + U_1)^2); \\ \delta_{\pm 4} &= 0, & (\text{if } \tau \neq 0). \end{aligned} \right\} \dots \dots \dots (18)$$

The eight asymptotes are therefore

$$\left. \begin{aligned} \omega &= U_1 l \pm \Omega_1, & (\tau \neq 0); \\ \omega &= l \text{ (two)}, & (\tau \neq (1 - U_1)^2); \\ \omega &= -l \text{ (two)}, & (\tau \neq (1 + U_1)^2); \\ \omega &= (U_1 \pm \sqrt{\tau})l, & (\tau \neq 0). \end{aligned} \right\} \dots\dots\dots (19)$$

In the special case where  $\tau$  is zero\* four of the asymptotes are parallel. Then

$$\sigma_{\pm 1} = \sigma_{\pm 4} = U_1^{-1}, \quad (\tau = 0)$$

and equations (16) and (17) are identities.  $\delta$  must now be determined from the relation

$$\varphi_6 \delta^4 + \varphi_5 \delta^3 + \varphi_4 \delta^2 + \varphi_3 \delta + \varphi_2 = 0 \quad \dots\dots\dots (20)$$

This yields the values

$$\left. \begin{aligned} \delta_{\pm 1} &= \pm (\sqrt{2} U_1)^{-1} \{ \Omega_0^2 + 1 + \sqrt{\Omega_0^4 + 2(\Omega_T^2 - \Omega_1^2) + 1} \}^{\frac{1}{2}}, \\ \delta_{\pm 4} &= \pm (\sqrt{2} U_1)^{-1} \{ \Omega_0^2 + 1 - \sqrt{\Omega_0^4 + 2(\Omega_T^2 - \Omega_1^2) + 1} \}^{\frac{1}{2}}. \end{aligned} \right\} \dots\dots\dots (21)$$

The points of intersection with the  $\omega$ -axis are given by the roots of

$$\omega^2 P_1(\omega^2, 0) = 0,$$

namely,

$$\left. \begin{aligned} \omega_{\pm 1}^P &= 0, \\ \omega_{\pm 2}^P &= \pm \frac{1}{2} \{ \sqrt{\Omega_0^2 + 4} - |\Omega_0| \}, \\ \omega_{\pm 3}^P &= \pm 1, \\ \omega_{\pm 4}^P &= \pm \frac{1}{2} \{ \sqrt{\Omega_0^2 + 4} + |\Omega_0| \}. \end{aligned} \right\} \dots\dots\dots (22)$$

The curve thus cuts the  $\omega$ -axis in eight real points, which are independent of  $U_0$  and  $\tau$ , and of which only four depend on  $|\Omega_0|$ .†

The intersections with the  $l$ -axis are given by

$$l^2 = 0 \quad \dots\dots\dots (23a)$$

and the six roots of the cubic in  $l^2$

$$\begin{aligned} (U_1^2 - \tau)l^6 + [2U_1^2 - \tau(2 - \Omega_1^2/U_1^2) - \Omega_0^2 - 1]l^4 \\ + [U_1^2 - \tau - 2 - \Omega_T^2 + (\Omega_1^2/U_1^2) + 2(\Omega_1/U_1)(\tilde{\Omega}_T/U_T)]l^2 - 1 = 0 \quad \dots\dots\dots (23b) \end{aligned}$$

### (b) The Standard Cases

Special cases often occur in which the expression for  $P(\omega, l)$  has factors. Such standard cases are conveniently designated by  $Cmn$ , where  $m$  (or  $n$ ) takes the values 0, 1, 2, 3, or  $T$  as the drift velocity (or static magnetic field) is respectively zero, parallel to the  $x$ -,  $y$ -, or  $z$ -axis or transverse to the  $x$ -axis.

In the case  $C11$

$$P(\omega, l) = X_0(Y_0 + \Omega_1 Z)(Y_0 - \Omega_1 Z) \quad \dots\dots\dots (24)$$

\* The other special cases,  $\tau = (1 - U_1)^2$ , are neglected on physical grounds. By virtue of its definition  $\tau < 1/3$ . The positive sign is therefore impossible and the negative sign would require  $U_1 > 0.423$ , necessitating a relativistic treatment.

† Thus when  $v=0$  the frequencies for which the refractive index vanishes are independent of  $U_0$  and  $\tau$ . On the other hand, as may easily be shown, the frequencies for which the group velocity is zero are, in general, functions of  $U_0$  and  $\tau$ .



and in the case  $CmT$

$$P(\omega, l) = RY_0[X_0(Z+1) - \Omega_T^2 Z] \dots \dots \dots (25)$$

In both these cases it is necessary to solve equations of only the first or second degree to plot the curve  $P=0$ .

The curves for a number of such standard cases have been constructed. Examples of the cases  $C11$  and  $C1T$  are illustrated by Figures 1 and 2.

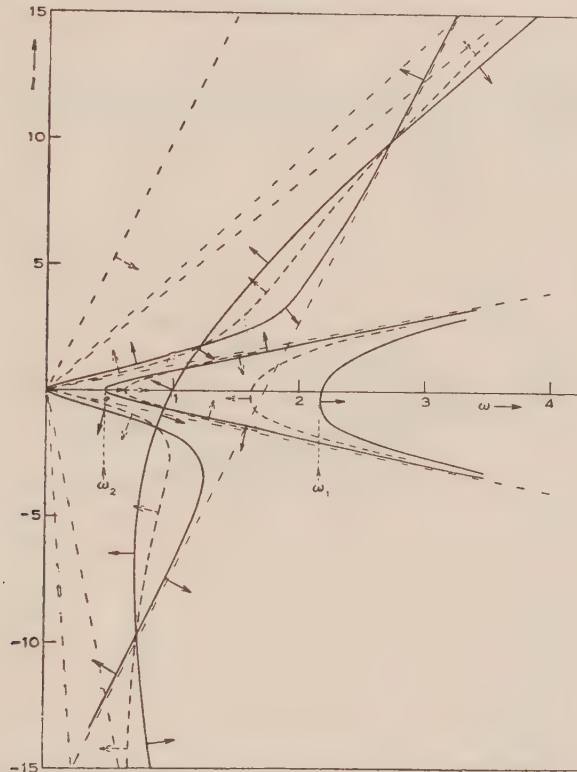


Fig. 1.—The wave number  $l$  as a function of the frequency  $\omega$  for the case  $C11$ .

$$U_1=0.1, U_T=0, \Omega_1=1.73, \Omega_T=0; \tau=0.02.$$

The units of velocity and frequency are taken as equal to  $c$  and  $p$  respectively.

$$\omega_1, \omega_2 = \frac{1}{2} \{ \sqrt{\Omega_1^2 + 4} \pm |\Omega_1| \}.$$

The general case is illustrated by Figures 3A and 3B, which have been computed and drawn by Mr. W. Moriarty. It is interesting to compare the special cases  $C11$  and  $C1T$  (Figs. 1 and 2) with this general case. When the magnetic field is entirely transverse (Fig. 2) the curves passing through the origin degenerate to a pair of straight lines, the remainder of the curves being substantially unaltered. When the field is entirely longitudinal (Fig. 1) there is another type of degeneracy. From Figure 3B we see that the curve intersecting the  $\omega$ -axis in  $\omega=1$  is asymptotic to  $\omega=\pm l$ . Hence in case  $C11$  the curves intersecting the axis in  $\omega=1$  and  $\omega=\omega_2$  must be regarded as having

discontinuous tangents at their two common points, if these curves are to conform with the general case in Figure 3.

This behaviour has previously been studied in the Magneto-Ionic Theory of propagation which is here represented by the standard case *CO*n with  $\tau=0$ . The  $(\omega, l)$  curves are then similar to the curves in Figures 3A and 3B but are symmetrical about the  $\omega$ -axis; also the asymptotes other than  $\omega = \pm l$  are now all vertical and displaced from the origin.

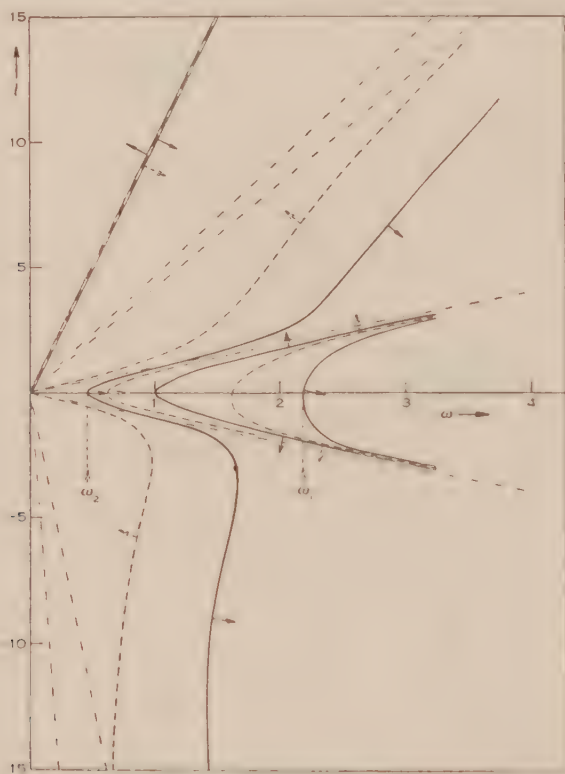


Fig. 2.—The wave number  $l$  as a function of the frequency  $\omega$  for the case C17.

$$U_1=0.1, U_T=0, \Omega_1=0, \Omega_T=1.73; \tau=0.02.$$

The units of velocity and frequency are taken as equal to  $c$  and  $p$  respectively.

$$\omega_1, \omega_2 = \frac{1}{2}(\sqrt{\Omega_T^2 + 4} \pm |\Omega_T|).$$

### (c) Waves of the First and Second Species

From the *P* curve we may obtain the first approximations, namely the roots of equation (6). In the case illustrated it is seen that for A2 waves (frequency given real), there are eight real solutions for  $l$  at the higher frequencies. In certain frequency bands, however, some of the eight solutions are complex.

In all the cases illustrated there are eight real values of the frequency ( $\omega$ ) for any given wavelength (A3 waves).

When these first approximations (roots of (6)) are real, the waves are called *waves of the First Species (FS)*. In the absence of collisions there is no net

transfer of energy to or from an *FS* wave. If, however, the first approximations are complex (and the real part is non-zero), the waves are termed waves of the *Second Species* (*SS*), and even in the absence of collisions, there is a net exchange of energy between these waves and the stream of electrons.

#### IV. WAVES OF THE FIRST SPECIES

In the notation of Section II,

$$l = l_0 - i\eta$$

for *A2* waves, and

$$\omega = \omega_0 - i\xi$$

for *A3* waves. For waves of the first species,  $l_0$ ,  $\eta$ ,  $\omega_0$ , and  $\xi$  are all real. Thus for *A2* waves the wave number and the coefficient of attenuation are respectively

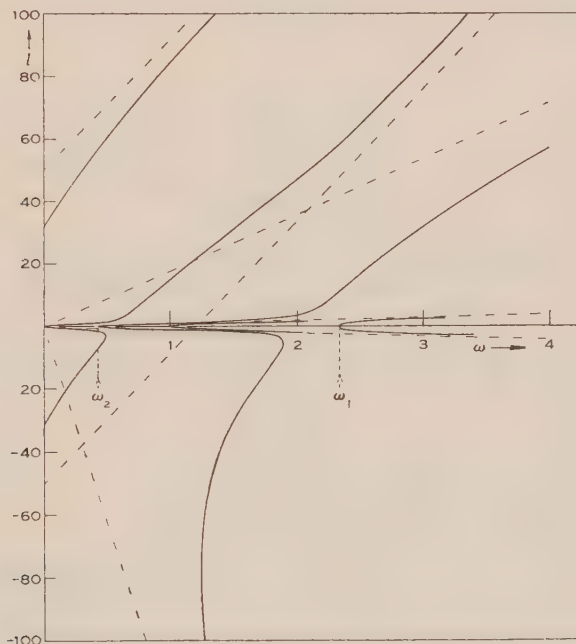


Fig. 3A.—The wave number  $l$  as a function of the frequency  $\omega$  for the general case *Cmn*.

$$U_1 = 0.024, \quad \Omega_1 = 1.2, \quad \tau = 0.001,$$

$$U_T = 0.018, \quad \Omega_T = 1.6, \quad \tilde{\Omega}_T \cdot U_T = 0.020$$

The units of velocity and frequency are taken as equal to  $c$  and  $p$  respectively.

$$\omega_1, \omega_2 = \frac{1}{2} \{ \sqrt{\Omega_0^2 + 4} \pm |\Omega_0| \}$$

$l_0$  and  $\eta$ , while for *A3* waves  $\omega_0$  and  $-\xi$  denote the frequency and coefficient of damping. Furthermore, for first species waves, the phase velocity  $Re(\omega)/Re(l)$  is given by the value of  $\sigma^{-1}$  on the *P* curve, whilst the group velocity  $dRe(\omega)/dRe(l)$  is the inverse of the slope of the *P* curve. Hence:

$$\left. \begin{array}{l} \text{An A2 FS wave is attenuated or spatially growing as} \\ \eta \sigma^{-1} \gtrless 0. \end{array} \right\} \dots (26)$$

An  $A3$   $FS$  wave is damped or secularly growing as  $\left. \begin{array}{l} \xi < 0. \\ \xi > 0. \end{array} \right\} \dots (27)$

The signs of these terms may conveniently be determined by the following method :

The  $P=0$  and  $Q=0$  curves (respectively continuous and broken) are drawn on the same diagram. Normal to these curves short arrows are drawn in the direction in which the function ( $P$  or  $Q$ ) increases. The sign of  $Q$  and

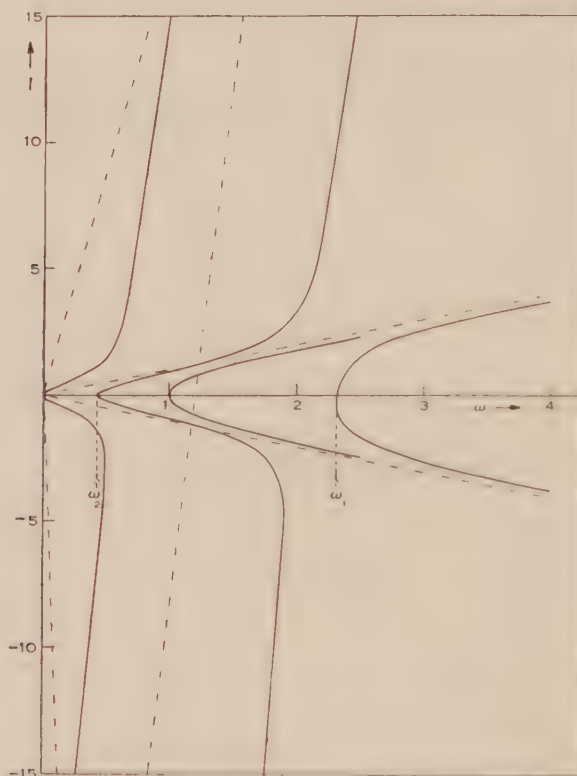


Fig. 3B.—Detail of Figure 3A near the  $\omega$ -axis.

$\sigma^{-1}$  at any point on the  $P$  curve is then given by inspection and the signs of  $\partial P/\partial \omega$ , and  $\partial P/\partial l$  may be found by the following rules :

$P'(\omega) > 0$  as the arrow on the  $P$  curve points in the direction of increasing or decreasing  $\omega$ .

$P'(l) > 0$  as the arrow on the  $P$  curve points in the direction of increasing or decreasing  $l$ .

(a) The Curve  $Q=0$

With the same transformation as that used for the curves  $P=0$ , we obtain the following expressions :

$$Q(\omega, l) = \omega^3 r Q_1(\omega^2, \sigma), \dots (28)$$



where

$$\left. \begin{aligned} Q_1(\omega^2, \sigma) &= \chi_4 \omega^4 + \chi_3 \omega^2 + \chi_0, \\ \chi_4 &= \zeta^2 (2\tau\sigma^2 - 3r^2), \\ \chi_3 &= \zeta [(2 + \Omega_0^2)\zeta - 4r^2 + 2\tau\sigma^2], \\ \chi_0 &= 2\zeta - r^2. \end{aligned} \right\} \dots (29)$$

In this case the values of  $\omega$  for any selected  $\sigma$  are given by the solutions of a quadratic. The  $Q$  curve is thus easily plotted.

The asymptotes to the curve are given by

$$l = \rho_n \omega + \gamma_n, \quad (n=0, \pm 1, \pm 2, \pm 3) \dots (30)$$

where

$$\left. \begin{aligned} \rho_0 &= U_1^{-1}, \\ \rho_{\pm 1} &= 1, \\ \rho_{\pm 2} &= -1, \\ \rho_{\pm 3} &= (U_1 \pm \sqrt{2\tau/3})^{-1}, \end{aligned} \right\} \dots (31)$$

and

$$\left. \begin{aligned} \gamma_0 &= 0, \\ \gamma_{\pm 1} &= 0, \quad \left( \tau \neq \frac{3}{2}(1 - U_1^2) \right); \\ \gamma_{\pm 2} &= 0, \quad \left( \tau \neq \frac{3}{2}(1 + U_1^2) \right); \\ \gamma_{\pm 3} &= 0, \quad (\tau \neq 0). \end{aligned} \right\} \dots (32)$$

We have also the special case

$$\tau = 0, \quad \gamma_{\pm 3} = (\sqrt{3}U_1)^{-1} \sqrt{\Omega_0^2 + 2} \dots (33)$$

The intercepts on the  $\omega$ -axis are given by the roots of

$$\omega^3 Q_1(\omega^2, 0) = 0,$$

namely

$$\left. \begin{aligned} \omega_0^Q &= \omega_{\pm 1}^Q = 0, \\ \omega_{\pm 2}^Q &= \pm (1/\sqrt{12})(\sqrt{\Omega_0^2 + 12} - |\Omega_0|), \\ \omega_{\pm 3}^Q &= \pm (1/\sqrt{12})(\sqrt{\Omega_0^2 + 12} + |\Omega_0|). \end{aligned} \right\} \dots (34)$$

The intercepts on the  $l$ -axis are given by

$$l^3 = 0 \dots (35a)$$

and the four roots of the quadratic in  $l^2$

$$(\tau - \frac{3}{2}U_1^2)l^4 + (1 + \frac{1}{2}\Omega_0^2 + \tau - 2U_1^2)l^2 + 1 - \frac{1}{2}U_1^2 = 0. \dots (35b)$$

### (b) The Arrows on the P and Q Curves

The points of intersection of the  $P$  and  $Q$  curves with the  $\omega$ -axis satisfy the relations

$$|\omega_{\pm 4}^P| \geq |\omega_{\pm 3}^Q| \geq |\omega_{\pm 3}^P| \geq |\omega_{\pm 2}^Q| \geq \omega_{\pm 2}^P \geq \omega_{\pm 1}^Q = \omega_{\pm 1}^P = \omega_0^Q \dots (36)$$

Consideration of  $P(\omega_n^P + \delta\omega, 0)$ , ( $\delta\omega > 0$ ) shows that it has the sign  $(-1)^n$ . Hence for the branches of  $P=0$  through  $\omega_n^P$ ,  $n=-3, -1, 2, 4$  we draw the arrows to the right and for the branches through  $\omega_n^P$ ,  $n=-4, -2, 1, 3$  to the left. Similarly the sign of  $Q(\omega_n^Q + \delta\omega, 0)$ , ( $\delta\omega > 0$ ) is  $(-1)^n$ , so that for branches through  $\omega_n^Q$ ,  $n=0, \pm 2$  we draw the arrows to the right and for those through  $\omega_n^Q$ ,  $n=\pm 1, \pm 3$  we draw them to the left.

It should be noted that this analysis determines the directions of the arrows only near the  $\omega$ -axis. If at any point the curves cross, it is necessary to make another examination of the signs of the functions.

(c) *Examples and Some Conclusions about FS Waves*

The curves  $Q=0$  and the arrows have been included in the examples of Figures 1 and 2. Using the method of Section IV we conclude as follows:

1. For the cases illustrated, there are certain frequency bands in which *FS A2* waves exhibit negative attenuation. (It should, however, be noted that as shown below in Section IV (d) the corresponding groups do not grow.)

2. For all the cases illustrated and for all other cases investigated, *FS A3* waves (waves of a given wavelength) are damped in time.

(d) *Groups of FS Waves*

As infinite wave trains have little physical significance, we shall here give some consideration to wave groups. Consider a narrow group of *FS A2* waves centred about the real frequency  $\omega$ , for which  $l=l_0-i\gamma$  ( $l_0, \gamma$  real). It may be shown that such a group moves with the velocity

$$\left. \begin{aligned} g &= \frac{d\omega}{dl_0} = - \frac{\partial P(\omega, l_0)/\partial l_0}{\partial P(\omega, l_0)/\partial \omega_0}, \\ &= \text{inverse of the slope of the } P \text{ curve.} \end{aligned} \right\} \dots (37)$$

$$\left. \begin{aligned} &\text{The group grows or diminishes according as} \\ &\qquad \qquad \qquad \eta g \leq 0. \end{aligned} \right\} \dots \dots \dots (38)$$

The sign of  $g$  may be read directly from the graph and the sign of  $\gamma$  determined by the method indicated above (Section IV). It is thus a relatively simple matter to determine the behaviour of groups of *FS A2* waves. The process may, however, be simplified by using the following theorem:

A group of *FS A2* waves centred about the frequency  $\omega$  and wave number  $l_0$  grows or diminishes with time according as the *FS A3* wave of wave number  $l_0$  grows or diminishes.

From the definitions of  $\eta$  and  $g$  (equations (8) and (37)),

$$\eta g = -\nu \frac{Q(\omega, l_0)}{\partial P(\omega, l_0)/\partial \omega_0},$$

which by equation (10) is equal to  $-\xi$ , where  $-\xi$  is the damping coefficient of an *FS A3* wave, of wave number  $l_0$ . Comparison of the conditions (27) and (38) then establishes the theorem.

Examples of growing *FS A2* waves may be found in Figures 1 and 2, but in all the examples so far studied no growing *FS A2* wave groups have been found.

## V. WAVES OF THE SECOND SPECIES

As the coefficients in  $P$  are real, complex solutions of  $P=0$  must occur in conjugate pairs. We thus have the following important conclusions:

1. If, for a given real frequency  $\omega$ , the roots of  $P=0$  are complex of the form

$$l_{10}, l_{20} = \alpha \pm i\beta, \quad (\alpha \neq 0), \dots\dots\dots (39)$$

a narrow wave group centred about  $\omega$  will (in the absence of collisions) grow in amplitude as it progresses.

2. If, for a given real wave number  $l$ , the roots of  $P=0$  are complex of the form

$$\omega_{10}, \omega_{20} = \gamma \pm i\delta, \quad (\gamma \neq 0), \dots\dots\dots (40)$$

one of the corresponding waves will (in the absence of collisions) everywhere grow in amplitude with the lapse of time.

When further approximations are made to take account of the collisions, these conclusions remain valid provided the collision frequency is less than a certain value.

TABLE 1  
SOME EXAMPLES OF  $A2$   $SS$  GROWING WAVES

Conditions		Frequency (in units $p$ )	Wavelength $2\pi/\alpha$ (in units $c/p$ )	Coeff <sup>t</sup> of Negative At- tenuation $\beta$ (in units $p/c$ )	Complex Refractive Index $l/\omega$
$C11$ $\Omega_1=1.73$	$U_1=0.1$	0.5	1.26	7.07	$-10 \pm i 14.1$
	$\tau=0.02$	1.26	1.94	0.78	$-2.58 \pm i 0.62$
$C1T$ $\Omega_T=1.73$	$U_1=0.1$				
	$\tau=0.02$	1.26	0.50	9.2	$-9.9 \pm i 7.4$

The real roots (yielding  $FS$  waves), and the regions in which complex roots (yielding  $SS$  waves) occur, are directly ascertained from the curves  $P=0$ . In frequency-bands where  $SS$  waves occur, these roots may be determined from the real roots and the relations between the coefficients in  $P$ . If there is only one pair of complex roots, this process yields only a quadratic equation. It is then most convenient to use the relations

– Sum of the roots = coefficient of second highest power,

Product of the roots = constant term.

Examples of  $A2$  growing waves obtained by means of the curves in Figures 1 and 2 are given in Table 1. The corresponding refractive indices are shown in the last column.

In the examples illustrated in Figures 1, 2, and 3,  $\tau > U_1^2$ . Under these conditions the curves which cut the  $\omega$ -axis in  $\omega = \omega_2$  do not cut the  $l$ -axis. When  $\tau < U_1^2$  these curves cut this axis at two points, and for low frequencies there is one pair less of  $SS$  waves.

From the form of the curves in the general case we may conclude that A2 SS growing waves can occur only for frequencies below

$$\omega_1 \equiv \omega_4^P = \frac{1}{2} [\sqrt{\Omega_0^2 + 4} + |\Omega_0|].$$

In all the cases investigated so far, no A3 SS waves (which grow in time) have been found, when motion of the positive ions is neglected. For with any real value of  $l$  the eight roots of  $P=0$  are all real. No proof has been found that this holds in general but it can be demonstrated in a number of special cases.

When the motions of the positive ions are also considered in the theory, A3 SS waves are found to exist under suitable conditions (Section VI).

When considering the application of the theory to non-uniform media, it is desirable to study the variation of the properties of the medium with the electron density  $N_0$ . It is then more convenient to plot curves of the refractive index  $M$  against

$$\xi = p^2/\omega^2 = (4\pi e^2/m\omega^2)N_0;$$

with this notation equation (6) becomes

$$\psi_3(M)\xi^3 + \psi_2(M)\xi^2 + \psi_1(M)\xi + \psi_0(M) = 0, \dots \dots \dots (12a)$$

where

$$\begin{aligned} \psi_3(M) &= r^2, \\ \psi_2(M) &= r^2[2\xi + \tau_0 M^2 - r^2], \\ \psi_1(M) &= \zeta[(r^2 - \gamma_{1T}^2)\zeta + r^2(\gamma_{1T}^2 - 2r^2 + 2\tau_0 M^2) + 2r\zeta M\gamma_{1T}(\tilde{\gamma}_T \cdot \mathbf{V}_T)], \\ \psi_0(M) &= \zeta^2[r^2\gamma_{1T}^2 - (r^2 - \tau_0 M^2)(r^2 - \gamma_{1T}^2)], \\ \mathbf{V}_0 &= \mathbf{U}_0/c, \quad \tilde{\gamma}_0 = \tilde{\Omega}_0/\omega, \quad \tau_0 = \tau/c^2, \\ r &= 1 - MV_1, \quad \zeta = M^2 - 1. \end{aligned}$$

As equation (12a) is a cubic in  $\xi$ , the  $(M, \xi)$  curve may easily be plotted by solving (12a) for selected values of  $M$ . These curves are appropriate for the discussion of Aspect 1, in which  $M$  is given real.

## VI. MOTION OF THE POSITIVE IONS INCLUDED

When the motions of the positive ions are also considered (3), the equation of dispersion is of the twelfth degree in  $\omega$  and  $l$ . Solutions of this equation in the general case have not as yet been considered. However, in the standard case in which the static electric and magnetic fields are both parallel to the direction of propagation, the equation may be factorized, yielding the three quartics

$$X X_1 - p^2 p_i^2 = 0, \dots \dots \dots (41)$$

$$Z(R + \Omega_1 - i\nu)(R_1 + \Omega_{11} - i\nu_1) + p^2 R(R_1 + \Omega_{11} - i\nu_1) + p_i^2 R_1(R + \Omega_1 - i\nu) = 0, \dots \dots \dots (42)$$

and

$$Z(R - \Omega_1 - i\nu)(R_1 - \Omega_{11} - i\nu_1) + p^2 R(R_1 - \Omega_{11} - i\nu_1) + p_i^2 R_1(R - \Omega_1 - i\nu) = 0, \dots \dots \dots (43)$$

where

$$\left. \begin{aligned} X_i &= R_i^2 - \tau_i l^2 - p_i^2 - i\nu_i R_i, \\ R_i &= \omega - U_{i1} l, \end{aligned} \right\} \dots \dots \dots (44)$$



$$\left. \begin{aligned} U_{i0} &= \text{mean drift velocity of the positive ions,} \\ \bar{\Omega}_{i0} &= eH_0/m_i c = \text{ion gyro-frequency vector,} \\ \nu, \nu_i &= \text{frequencies of collision, with gas molecules, of an} \\ &\quad \text{electron and ion respectively,} \\ \tau_i &= \text{one-third of the mean square velocity of agitation} \\ &\quad \text{of the ions,} \\ p_i &= \sqrt{4\pi N_0 e^2 / m_i}, \\ m_i &= \text{mass of a positive ion,} \end{aligned} \right\} \dots (45)$$

the mean electron and ion densities are assumed equal, and the ions are assumed to carry a charge  $-e$ .

Equation (41) is independent of the magnetic field.

When  $\nu, \nu_i$  are small we may use the same method of approximation as that in Section II, taking as first approximations the solutions of these equations with  $\nu = \nu_i = 0$ . Applying the usual change of variable, (41) then reduces to the quadratic

$$\omega^2 = \frac{p^2}{r^2 - \tau \sigma^2} + \frac{p_i^2}{r_i^2 - \tau_i \sigma^2}, \dots (46)$$

and (42), (43) to the cubics

$$\psi_3 \omega^3 \pm \psi_2 \omega^2 + \psi_1 \omega \pm \psi_0 = 0, \dots (47)$$

where

$$\left. \begin{aligned} \psi_3(\sigma) &= r r_i \zeta, \\ \psi_2(\sigma) &= (r \Omega_{i1} + r_i \Omega_1) \zeta, \\ \psi_1(\sigma) &= r r_i (p^2 + p_i^2) + \Omega_1 \Omega_{i1} \zeta, \\ \psi_0(\sigma) &= p^2 r \Omega_{i1} + p_i^2 r_i \Omega_1, \end{aligned} \right\} \dots (48)$$

and

$$\left. \begin{aligned} r &= 1 - U_1 \sigma \\ r_i &= 1 - U_{i1} \sigma \end{aligned} \right\} \dots (49)$$

The asymptotes are

$$\left. \begin{aligned} \omega &= U_1 l \pm \Omega_1, \\ \omega &= U_{i1} l \pm \Omega_{i1}, \\ \omega &= l, \text{ (two)} \\ \omega &= -l, \text{ (two)} \\ \omega &= (U_1 \pm \sqrt{\tau}) l, \quad (\tau \neq 0) \\ \omega &= (U_{i1} \pm \sqrt{\tau_i}) l, \quad (\tau_i \neq 0). \end{aligned} \right\} \dots (50)$$

If  $\tau = 0$  the fifth pair of asymptotes becomes

$$\omega = U_1 l \pm p, \dots (51a)$$

and if  $\tau_i = 0$  the sixth pair becomes

$$\omega = U_{i1} l \pm p_i \dots (51b)$$

As an example Figure 4 shows the  $(\omega, l)$  curve corresponding to equation (46), with the following values for the constants:

$$U_1 = 0.1, \quad U_{i1} = -0.01, \quad p_i = 0.1, \quad \tau = 0, \quad \tau_i = 0,$$

where we have again adopted  $p$  and  $c$  as units of frequency and velocity respectively. It is evident from this figure that for low frequencies growing A2 SS waves again occur. Examples of such A2 growing waves have been given in an earlier publication(3).

Figure 4 also demonstrates the existence of A3 SS waves, i.e. waves of a given real wavelength which everywhere grow with the lapse of time. For when  $|l| < 12.2$  there are two real values for  $\omega$  and a complex conjugate pair of the form (40), namely,

$$\omega_{10}, \omega_{20} = \gamma \pm i\delta.$$

These complex roots may be determined by the method discussed in Section V. Values of  $\gamma$  and  $\delta$  are also shown in Figure 4.

The graphical method developed here for the study of electro-magneto-ionic waves may also be used to study the plane waves which arise from the interaction of two streams of electrons and which have been discussed recently



Fig. 4.—The wave number  $l$  as a function of the frequency  $\omega$  for four of the twelve possible waves in the case C11, when the motion of the positive ions is considered.

$$U_1 = 0.1, \quad U_T = 0, \quad \Omega_T = 0, \quad \tau = 0,$$

$$U_{H1} = 0.01, \quad U_{HT} = 0, \quad p_2 = 0.1, \quad \tau_1 = 0.$$

The units of velocity and frequency are taken as equal to  $c$  and  $p$  respectively.

The complex frequencies  $\omega_{01}, \omega_{02} = \gamma \pm i\delta$ , which occur in a certain wave number band, are shown dotted.

by Haefl(7, 8), Nergaard(9), and Pierce and Hebenstreit(10). The relevant dispersion equation is the same as (46) with  $\tau = \tau_1 = 0$  and  $U_0, p_1$  representing respectively the drift velocity and plasma frequency of the second stream of electrons. By this method it can be shown that under certain conditions in such multiple streams of electrons, waves can occur which at every place grow in time (A3 SS waves).

## VII. ACKNOWLEDGMENTS

We are indebted to Mr. W. Moriarty for his assistance in computing and drawing the curves in Figures 3A and 3B, and to Dr. J. L. Pawsey for assisting with the final reproduction of all the figures.

## VIII. REFERENCES

- (1) BAILEY, V. A.—*Nature* **161** : 599 (1948).
- (2) BAILEY, V. A.—*J. Roy. Soc. N.S.W.* **82** : 107 (1948).
- (3) BAILEY, V. A.—*Aust. J. Sci. Res. A* **1** : 351 (1948).
- (4) JAHNKE, E., and EMDE, F.—“Tables of Functions.” Addenda pp. 24-30. (Dover Publications : New York, 1945.)
- (5) EDWARDS, J.—“Differential Calculus.” p. 183. (McMillan & Co. : London, 1912.)
- (6) EDWARDS, J.—*Ibid.* p. 187, Method III.
- (7) HAEFF, A. V.—*Phys. Rev.* **74** : 1532 (1948).
- (8) HAEFF, A. V.—*Proc. Inst. Radio Engrs.* **37** : 4 (1949).
- (9) NERGAARD, L. S.—*R.C.A. Rev.* **9** : 585 (1948).
- (10) PIERCE, J. R., and HEBENSTREIT, W. B.—*Bell Syst. Tech. J.* **28** : 33 (1949).

# TRANSIENTS IN AN IONIZED MEDIUM WITH APPLICATIONS TO BURSTS OF SOLAR NOISE

By J. C. JAEGER\* and K. C. WESTFOLD†

[Manuscript received February 7, 1949]

## Summary

Exact solutions of a number of transient problems on linear propagation in a homogeneous ionized medium without magnetic field are given. The Fourier transforms of the solutions are discussed and their importance in this type of work emphasized. The effect of inhomogeneity of the medium is studied, and it is shown that if a localized disturbance takes place in the solar atmosphere at a level at which the collision frequency is  $\nu$  and the frequency of plasma oscillations is  $\omega_0$ , radiation will be emitted on all frequencies greater than  $\omega_0$  with intensity determined by the nature of the disturbance and with damping determined by  $\nu$ . There is a small difference between the times of arrival on different frequencies, and on each frequency there is a direct wave and an echo wave which arrives some seconds later. Many of the phenomena of bursts of solar noise are consistent with these predictions.

## I. INTRODUCTION

The properties of steady periodic electromagnetic radiation in ionized media have been extensively studied, but little attention has been paid to transient phenomena in them. We give here some exact solutions for simple cases in order to show the general nature of the transient behaviour of such media, and the way in which it differs from the steady state behaviour: this difference is found to be conspicuous, and possibly of importance in the interpretation of observations on solar noise.

In order to obtain explicit solutions which show the general nature of the phenomena clearly, we shall discuss in detail the propagation of transients in one dimension in a homogeneous ionized medium without magnetic field. This is done in Sections II and III, neglecting collisions, and in Section IV including the effect of collisions. The Fourier transforms of the solutions are discussed in Section V. The effects of a magnetic field, and the possibilities of extension to two- or three-dimensional systems, are discussed in Section VI. The effects of inhomogeneity of the medium which are fundamental for application to the solar atmosphere are studied in Section VII. Finally, in Sections VII and VIII the physical consequences of the results are studied and compared with those to be expected on steady state theory; also experimental results on bursts of solar noise are discussed with reference to them.

\* Department of Mathematics, University of Tasmania.

† Division of Radiophysics, C.S.I.R.O.



## II. GENERAL EQUATIONS

The equations of propagation in an ionized medium are

$$\text{curl } \mathbf{E} = -\mu_v \frac{\partial \mathbf{H}}{\partial t}, \quad \dots \quad (1)$$

$$\text{curl } \mathbf{H} = \frac{\partial}{\partial t}(\epsilon_v \mathbf{E} + \mathbf{P}), \quad \dots \quad (2)$$

$$\frac{\partial^2 \mathbf{P}}{\partial t^2} + \nu \frac{\partial \mathbf{P}}{\partial t} = \epsilon_v \omega_0^2 \mathbf{E} + \frac{e\mu_v}{m} \frac{\partial \mathbf{P}}{\partial t} \times \mathbf{H}, \quad \dots \quad (3)$$

where

$$\omega_0^2 = \frac{Ne^2}{\epsilon_v m} \quad \dots \quad (4)$$

In the above,  $N$  is the electron density,  $\nu$  the electron collision frequency,  $e$  and  $m$  the charge and mass of the electron,  $\mathbf{P}$  the polarization, and the other symbols have their usual meanings. Rationalized M.K.S. units are used.  $\omega_0/2\pi$  is often referred to as the frequency of plasma oscillations of the medium.

The problem which will be considered here is that of the propagation through the medium of a disturbance which is initiated at zero time,  $t=0$ . To obtain explicit results it is necessary to consider extremely simple systems, and accordingly we shall specialize (1)–(3) to the one-dimensional case in which all quantities are independent of  $y$  and  $z$ , and in which, further, there is no external magnetic field and the non-linear terms  $\mathbf{P} \times \mathbf{H}$  in (3) are neglected. In this case the equations break up into three sets of two involving the  $x$ -,  $y$ -, and  $z$ -components, respectively, of  $\mathbf{P}$  and  $\mathbf{E}$ . We shall consider the  $z$ -components only, and write  $P$  and  $E$  for  $P_z$  and  $E_z$ . Then the equations for these are

$$\frac{\partial^2 P}{\partial t^2} + \nu \frac{\partial P}{\partial t} = \epsilon_0 \omega_0^2 E, \quad \dots \quad (5)$$

$$\frac{\partial^2 E}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2} - \mu_v \frac{\partial^2 P}{\partial t^2} = 0, \quad \dots \quad (6)$$

where  $c^2 = 1/\epsilon_v \mu_v$ .

III. THE CASE  $\nu=0$ .  $N$  CONSTANT

Here (5) and (6) reduce to

$$\frac{\partial^2 E}{\partial x^2} - \frac{1}{c^2} \frac{\partial^2 E}{\partial t^2} - \frac{\omega_0^2}{c^2} E = 0 \quad \dots \quad (7)$$

This has to be solved with given initial values\*  $E_0(x)$  and  $E_1(x)$  of  $E$  and  $\partial E/\partial t$ . We use the Laplace transformation method with the notation of Carslaw and Jaeger(1), namely

$$\bar{E} = \int_0^\infty e^{-pt} E dt \quad \dots \quad (8)$$

for the Laplace transform of  $E$ . The subsidiary equation for  $\bar{E}$  is then

$$\frac{d^2 \bar{E}}{dx^2} - \frac{p^2 + \omega_0^2}{c^2} \bar{E} = -\frac{p}{c^2} E_0(x) - \frac{1}{c^2} E_1(x) \quad \dots \quad (9)$$

\* If the initial values of  $H_y$  and  $\partial P_z/\partial t$  are specified,  $E_1(x)$  may be replaced by the initial value of  $[\partial H_y/\partial x - \partial P_z/\partial t]/\epsilon_v$ .

This has to be solved with appropriate boundary conditions. We shall first consider the case of the infinite region  $-\infty < x < \infty$  in which the boundary conditions require that  $\bar{E}$  be bounded as  $x \rightarrow -\infty$ . To solve (9) with these conditions, we first write the Green's function  $G(x, \xi)$  for the homogeneous equation corresponding to (9); this is

$$G(x, \xi) = \begin{cases} -\frac{1}{2q} e^{-q(x-\xi)}, & x > \xi \\ -\frac{1}{2q} e^{q(x-\xi)}, & x < \xi \end{cases} \dots\dots\dots (10)$$

where 
$$q^2 = \frac{p^2 + \omega_0^2}{c^2} \dots\dots\dots (11)$$

The problem of (9) is then (cf. Carslaw and Jaeger(1), Section 92, for details of the procedure applied to a similar problem),

$$\begin{aligned} \bar{E} = & -\frac{1}{c^2} \int_{-\infty}^{\infty} G(\xi, x) \{pE_0(\xi) + E_1(\xi)\} d\xi \\ & - \frac{1}{2qc^2} \int_{-\infty}^x e^{q(\xi-x)} \{pE_0(\xi) + E_1(\xi)\} d\xi \\ & + \frac{1}{2qc^2} \int_x^{\infty} e^{-q(x-\xi)} \{pE_0(\xi) + E_1(\xi)\} d\xi \dots\dots\dots (12) \end{aligned}$$

The simplest results, and the easiest to interpret physically, are obtained by taking the arbitrary functions to be Dirac delta functions.

If we take  $E_1(x) = \delta(x)$ ,  $E_0(x) = 0$ , corresponding to an initial value of  $\partial E / \partial t$  which is very great at the origin and zero elsewhere, we find from (12), using the value (11) of  $q$ ,

$$\bar{E} = \frac{1}{2c(p^2 + \omega_0^2)^{\frac{1}{2}}} e^{-(x/c)\sqrt{(p^2 + \omega_0^2)}}, \quad x > 0 \dots\dots\dots (13)$$

It follows from a table of Laplace transforms that

$$E = \begin{cases} \frac{1}{2c} J_0 \left\{ \omega_0 \sqrt{\left(t^2 - \frac{x^2}{c^2}\right)} \right\}, & t > \frac{x}{c} \\ 0, & t < \frac{x}{c} \end{cases} \dots\dots\dots (14)$$

where  $J_0(x)$  is the Bessel function of the first kind of order zero.

The solution (14) corresponds to a wave front propagated with velocity  $c$ , followed by an infinite train of waves of slowly varying wavelength. The medium does not oscillate with the plasma frequency  $\omega_0/2\pi$ , but at a frequency which tends asymptotically to this value for large values of the time. This general behaviour appears in all cases.

If in (12) we take  $E_1(x) = 0$ ,  $E_0(x) = \delta(x)$ , corresponding to an initial value of  $E$  which is very great at the origin and zero elsewhere, we find

$$\bar{E} = \frac{p}{2c(p^2 + \omega_0^2)^{\frac{1}{2}}} e^{-(x/c)\sqrt{(p^2 + \omega_0^2)}} \dots\dots\dots (15)$$

$$E = \frac{1}{2c} \delta \left( t - \frac{x}{c} \right) - \frac{\omega_0 t J_1 [\omega_0 \sqrt{(t^2 - x^2/c^2)}]}{2c \sqrt{(t^2 - x^2/c^2)}}, \quad t > \frac{x}{c}$$

$$= 0, \quad t < \frac{x}{c} \quad \left. \vphantom{\frac{1}{2c}} \right\} \dots (16)$$

In the general case in which  $E_0(x)$  and  $E_1(x)$  may be arbitrary functions, we find from (12)

$$E = \frac{1}{2} \{ E_0(x-ct) + E_0(x+ct) \}$$

$$+ \frac{1}{2c} \int_{x-ct}^{x+ct} \left\{ E_1(\xi) J_0 \{ \omega_0 [t^2 - (x-\xi)^2/c^2]^{1/2} \} \right.$$

$$\left. - E_0(\xi) \frac{\omega_0 t J_1 [\omega_0 \{ t^2 - (x-\xi)^2/c^2 \}^{1/2}]}{[t^2 - (x-\xi)^2/c^2]^{1/2}} \right\} d\xi \dots (17)$$

This corresponds to disturbances, each of half the original waveform, moving in opposite directions, and each followed by a "tail".

Instead of considering the propagation of an initial disturbance, we may suppose  $E$  is prescribed at some point and study the disturbance radiated. Suppose, for example, that we consider the semi-infinite region  $x > 0$  with zero initial values of  $E$  and  $\partial E / \partial t$ , and that  $E$  is to have the value unity at  $x=0$  for  $t > 0$ . Then we have to solve (9) with  $E_0 = E_1 = 0$ , and with

$$\bar{E} = 1/p, \text{ when } x=0.$$

The solution is

$$\bar{E} = \frac{1}{p} e^{-(x/c) \sqrt{(p^2 + \omega_0^2)}}, \dots (18)$$

and

$$E = 1 - \frac{\omega_0 x}{c} \int_{x/c}^t \frac{J_1 \{ \omega_0 \sqrt{(\tau^2 - x^2/c^2)} \}}{\sqrt{(\tau^2 - x^2/c^2)}} d\tau, \quad t > x/c$$

$$= 0, \quad t < x/c \quad \left. \vphantom{\frac{\omega_0 x}{c}} \right\} \dots (19)$$

If  $E$  has the value

$$E = \sin \omega_1 t, \quad 0 < t < \pi/\omega_1$$

$$= 0, \quad t > \pi/\omega_1, \quad \left. \vphantom{\sin} \right\} \dots (20)$$

at  $x=0$ , that is, a single pulse of sinusoidal form, we have in place of (18)

$$\bar{E} = \frac{\omega_1}{p^2 + \omega_1^2} e^{-(x/c) \sqrt{(p^2 + \omega_0^2)}} [1 + e^{-\pi p/\omega_1}] \dots (21)$$

#### IV. THE CASE $\nu \neq 0$ . $N$ CONSTANT

We now consider how the above results are modified if collision damping is not neglected. Proceeding as in Section III, the subsidiary equations corresponding to (5) and (6) are

$$p(p + \nu) \bar{P} - \varepsilon_v \omega_0^2 \bar{E} = (p + \nu) P_0(x) + P_1(x)$$

$$\frac{d^2 \bar{E}}{dx^2} - \frac{p^2}{c^2} \bar{E} - \mu_v p^2 \bar{P} = -\frac{p}{c^2} E_0(x) - \frac{1}{c^2} E_1(x) - \mu_v p P_0(x) - \mu_v P_1(x),$$

where  $P_0(x)$  and  $P_1(x)$  are the initial values of  $P$  and  $\partial P / \partial t$ .

It follows that

$$\frac{d^2 E}{dx^2} - \left\{ p^2 + \frac{p\omega_0^2}{c^2(p+\nu)} \right\} E = -\frac{pE_0(x) + E_1(x)}{c^2} - \frac{\mu_\nu \nu P_1(x)}{p+\nu} \dots \quad (22)$$

As in Section III, the solution of this in  $-\infty < x < \infty$  is

$$E = \frac{1}{2q} \int_{-\infty}^x e^{q(\xi-x)} \left\{ \frac{pE_0(\xi) + E_1(\xi)}{c^2} + \frac{\mu_\nu \nu P_1(\xi)}{p+\nu} \right\} d\xi \\ + \frac{1}{2q} \int_x^{\infty} e^{q(x-\xi)} \left\{ \frac{pE_0(\xi) + E_1(\xi)}{c^2} + \frac{\mu_\nu \nu P_1(\xi)}{p+\nu} \right\} d\xi \dots \quad (23)$$

where, now,

$$q^2 = \frac{p^2}{c^2} + \frac{\omega_0^2 p}{c^2(p+\nu)} \dots \quad (24)$$

If we consider the case  $E_0(x) = P_1(x) = 0$ ,  $E_1(x) = \delta(x)$ , which led to the simplest result (14) in Section III, we find from (23)

$$E = \frac{1}{2c} \left\{ p^2 + \frac{p\omega_0^2}{p+\nu} \right\}^{-\frac{1}{2}} \exp \left\{ -\frac{x}{c} \left[ p^2 + \frac{p\omega_0^2}{p+\nu} \right]^{\frac{1}{2}} \right\} \dots \quad (25)$$

Even in this case there is no simple expression for  $E(t)$  in terms of known functions. It is easy to show that, as before, the wave front is propagated with velocity  $c$ . The only other information which may be obtained easily is the value of  $E$  at  $x=0$ , the point where the disturbance originated. Putting  $x=0$  in (25) and writing

$$p' = p + \frac{1}{2}\nu, \quad \omega_0'^2 = \omega_0^2 - \frac{1}{4}\nu^2,$$

(25) gives

$$E = \frac{1}{2c(p'^2 - \frac{1}{4}\nu^2)^{\frac{1}{2}} (p'^2 + \omega_0'^2)^{\frac{1}{2}}} \\ = \frac{1}{2c} \left( 1 + \frac{\nu}{2p'} + \dots \right) \frac{1}{(p'^2 + \omega_0'^2)^{\frac{1}{2}}}.$$

It follows that

$$E = \frac{1}{2c} e^{-\frac{1}{2}\nu t} J_0(\omega_0' t) + \frac{\nu}{4c\omega_0'} e^{-\frac{1}{2}\nu t} \int_0^{\omega_0' t} J_0(\tau) d\tau + \dots \quad (26)$$

If  $\nu/\omega_0$  is small,  $\omega_0'$  is indistinguishable from  $\omega_0$ , and the dominating term of (26) is just the solution (14) for the case  $\nu=0$  with the damping factor  $\exp(-\frac{1}{2}\nu t)$ . It is thus reasonable to suggest, though it is by no means proved, that a damping factor of this type will appear in the emitted wave train.

## V. THE FOURIER SPECTRUM OF THE RADIATION

If  $\bar{E}(p)$  is the Laplace transform of  $E(t)$ , and the Fourier transform of  $E(t)$  exists, this latter is given by

$$\bar{E}(i\omega) \dots \quad (27)$$

It follows (Carslaw and Jaeger(1), Section 119) that the amplitude of the component of frequency  $\omega/2\pi$  in the radiation is proportional to

$$|\bar{E}(i\omega)| \dots \quad (28)$$



This quantity (28) is much easier to evaluate than the actual radiation  $E(t)$  and it gives a great deal of useful information.

Considering first the case leading to (13), we find from (13)

$$\left. \begin{aligned} |\bar{E}(i\omega)| &= \frac{1}{2c\sqrt{(\omega^2 - \omega_0^2)}}, & \omega > \omega_0 \\ |\bar{E}(i\omega)| &= \frac{1}{2c\sqrt{(\omega_0^2 - \omega^2)}} e^{-(x/c)\sqrt{(\omega_0^2 - \omega^2)}}, & \omega < \omega_0. \end{aligned} \right\} \dots (29)$$

It appears from (29) that frequencies less than  $\omega_0/2\pi$  are attenuated rapidly as they pass through the medium so they will not appear in the emitted radiation. For frequencies higher than  $\omega_0/2\pi$  the amplitude falls off like  $(\omega^2 - \omega_0^2)^{-\frac{1}{2}}$ , that is, like  $(1/\omega)$  at very high frequencies.

The effect of collisions is mainly to round off the infinity of (29) at  $\omega = \omega_0$ . To see this we consider the result (25) which corresponds to (13) except that collisions are not neglected. It follows from (25) that

$$\bar{E}(i\omega) = \frac{1}{2c} \left\{ \frac{\omega_0^2}{1 - i\nu/\omega} - \omega^2 \right\}^{-\frac{1}{2}} \exp \left\{ -\frac{x}{c} \left[ \frac{\omega_0^2}{1 - i\nu/\omega} - \omega^2 \right]^{\frac{1}{2}} \right\} \dots (30).$$

If  $\omega < \omega_0$  the exponential term in (30) again causes  $|\bar{E}(i\omega)|$  to be attenuated rapidly with distance, so that frequencies less than  $\omega_0/2\pi$  do not appear. If  $\omega > \omega_0$  and we neglect  $(\nu/\omega)^2$ , the exponential term in  $|\bar{E}(i\omega)|$  becomes

$$\exp \left\{ -\frac{\nu x \omega_0^2}{2c\omega^2} \left( 1 - \frac{\omega_0^2}{\omega^2} \right)^{-\frac{1}{2}} \right\}, \dots (31)$$

this is just the attenuation given by the Lorentz theory for a wave of frequency  $\omega/2\pi$  travelling a distance  $x$  in the medium. If  $\omega = \omega_0$  and  $(\nu/\omega_0)^2$  is neglected, (31) is to be replaced by

$$\exp \{ -(x/c) [\frac{1}{2}\omega_0\nu]^{\frac{1}{2}} \}.$$

The other part of  $|\bar{E}(i\omega)|$  given by (30), which determines the amplitude of the component of frequency  $\omega/2\pi$  emitted, is

$$\frac{1}{2c} \left\{ \frac{1 + \nu^2/\omega^2}{(\omega^2 - \omega_0^2)^2 + \omega^2\nu^2} \right\}^{1/4}, \dots (32)$$

and, if  $\nu/\omega_0$  is small, this is very near to the value  $[2c(\omega^2 - \omega_0^2)^{\frac{1}{2}}]^{-1}$  of (29) except for values of  $\omega$  near  $\omega_0$ . When  $\omega = \omega_0$ , (32) has the value

$$\frac{(\nu^2 + \omega_0^2)^{1/4}}{2c\omega_0\nu^{\frac{1}{2}}}, \dots (33)$$

in place of the infinity in (29).

Thus the effect of collisions is to replace the infinity at  $\omega = \omega_0$  in the simple law (29) by a very sharp peak. It must be remembered that the peak actually observed in a receiver would be much less sharp than this, since (i) the observed effect will be an average of (32) over the bandwidth of the receiver, (ii) the attenuation corresponding to (31) in passing out to the receiver will be much greater for frequencies near  $\omega_0/2\pi$  than for the higher frequencies, and (iii) the

spectral components of the group move with different velocities so that the instantaneous spectral intensity distribution will vary along a trajectory.

We have discussed the simple case above in detail to show the general nature of the distribution with frequency of the energy radiated, and the effect of collisions on it. The formula for this distribution varies greatly with the nature of the original disturbance, but always preserves the features described above, namely that there is no radiation on frequencies less than  $\omega_0/2\pi$ , and that collisions have little effect except at frequencies near  $\omega_0/2\pi$ .

To illustrate the effect of the initial disturbance we give below the values of  $|\bar{E}(i\omega)|$  for  $\omega > \omega_0$  and  $v=0$  for some typical cases.

- (i)  $E_1(x)=\delta(x)$ ;  $E_0(x)=0$ ;  $\bar{E}$  given by (13);

$$|\bar{E}(i\omega)| = \frac{1}{2c\sqrt{(\omega^2 - \omega_0^2)}}, \quad \dots\dots\dots (34)$$

this is the case discussed above.

- (ii)  $E_1(x)=0$ ;  $E_0(x)=\delta(x)$ ;  $\bar{E}$  given by (15);

$$|\bar{E}(i\omega)| = \frac{\omega}{2c\sqrt{(\omega^2 - \omega_0^2)}}, \quad \dots\dots\dots (35)$$

- (iii)  $E_1(x)=1$ ,  $|x| < a$ ,  $E_1(x)=0$ ,  $|x| > a$ ;  $E_0(x)=0$ ;  $\bar{E}$  given by (12);

$$|\bar{E}(i\omega)| = \frac{1}{(\omega^2 - \omega_0^2)} \left| \sin \frac{a}{c} \sqrt{(\omega^2 - \omega_0^2)} \right|, \quad |x| > a \quad \dots (36)$$

- (iv)  $E_1(x)=0$ ;  $E_0(x)=1$ ,  $|x| < a$ ,  $E_0(x)=0$ ,  $|x| > a$ ;  $\bar{E}$  given by (12);

$$|\bar{E}(i\omega)| = \frac{\omega}{\omega^2 - \omega_0^2} \left| \sin \frac{a}{c} \sqrt{(\omega^2 - \omega_0^2)} \right|, \quad |x| > a \quad \dots (37)$$

- (v)  $E_0(x)=0$ ;  $E_1(x)=0$ ;  $E=1$ , when  $x=0$ ;  $\bar{E}$  given by (18);

$$|\bar{E}(i\omega)| = \frac{1}{\omega} \quad \dots\dots\dots (38)$$

- (vi)  $E_0(x)=0$ ;  $E_1(x)=0$ ;  $E=\delta(t)$ , when  $x=0$ ;

$$|\bar{E}(i\omega)| = 1 \quad \dots\dots\dots (39)$$

- (vii)  $E_0(x)=0$ ;  $E_1(x)=0$ ;  $E$  given at  $x=0$  by (20), and  $\bar{E}$  by (21);

$$|\bar{E}(i\omega)| = \left| \frac{2\omega_1}{\omega^2 - \omega_1^2} \right| \left| \cos \frac{\pi\omega}{2\omega_1} \right| \quad \dots\dots\dots (40)$$

- (viii)  $E_0(x)=0$ ;  $E_1(x)=0$ ;  $P_1(x)=\delta(x)$ ;  $\bar{E}$  given by (23);

$(v/\omega_0)^2$  negligible;

$$|\bar{E}(i\omega)| = \frac{\mu_v v c}{2\omega \sqrt{(\omega^2 - \omega_0^2)}} \quad \dots\dots\dots (41)$$

To draw qualitative conclusions from these it is simplest to consider their behaviour for large values of  $\omega$ . From (35), (34), and (41) it appears that for concentrated initial disturbances in  $E$ ,  $\partial E/\partial t$ , and  $\partial P/\partial t$  (the latter is an initial current),  $|\bar{E}(i\omega)|$  falls off like  $\omega^0$ ,  $\omega^{-1}$ , and  $\omega^{-2}$  respectively. If the initial disturbances are not concentrated but spread over a finite region,  $|\bar{E}(i\omega)|$  falls off more rapidly; for example, if  $a$  is small, (36) behaves like (34), but if

$a$  is large it falls off like  $\omega^{-2}$ . The same remark applies to (35) and (37). Similar results apply to the radiation produced by the application of a field at a point; if a constant electric field is suddenly applied,  $|\bar{E}(i\omega)|$  given by (38) falls off like  $\omega^{-1}$ ; for an impulsive applied electric field  $|\bar{E}(i\omega)|$  given by (39) is independent of frequency; while for a pulse of finite length  $|\bar{E}(i\omega)|$  given by (40) decreases like  $\omega^{-2}$ .

It appears that the amplitude of the component of frequency  $\omega/2\pi$  in the radiation may be expected to decrease with frequency according to some law such as  $\omega^{-1}$  or  $\omega^{-2}$  at high frequencies, and that the exact power in the law depends on the nature of the initial disturbance and whether it is very concentrated or spread over a relatively wide region. Thus if this law can be determined experimentally it may be possible to draw conclusions about the type of initial disturbance.

## VI. EXTENSION TO MORE COMPLICATED SYSTEMS

The preceding discussion has been confined to one-dimensional systems: it is possible to solve some similar problems in two or three dimensions without magnetic field, but the results are more complicated and the general behaviour of the solutions is not radically changed.

It is also possible to go some way in the discussion of one-dimensional propagation in the presence of an external magnetic field. It is known(2) that in the presence of such a field three principal directions appear in the medium, each associated with a characteristic frequency and damping of plasma oscillation, and that the phenomena associated with propagation in these directions are relatively simple to discuss. The theory of Sections II-V may be extended to propagation of transients in these directions with the results obtained above, except that now the characteristic plasma frequency and damping for the direction considered have to be used.

## VII. NON-HOMOGENEOUS MEDIA

In attempting to discuss the propagation of transients in the solar corona, we have to take into account the fact that it is non-homogeneous with an average electron density varying approximately inversely as a high power of the distance. While it is possible to obtain explicit solutions of (7) for some simple cases in which  $\omega_0^2$  is a function of  $x$ , these do not include anything like the actual law of variation in the corona, so we confine ourselves here to qualitative discussion.

We suppose, then, that a transient is produced by a concentrated initial disturbance at a level in the solar corona at which the frequency of plasma oscillation is  $\omega_0/2\pi$ ; the magnetic field being neglected as before. We conclude from the discussion above that:

- (i) radiation will be emitted on all frequencies greater than  $\omega_0/2\pi$  and there will be no radiation on frequencies less than  $\omega_0/2\pi$ ;
- (ii) the intensity of the radiation on any frequency  $\omega/2\pi$  will normally decrease with frequency, for large values of the frequency this decrease will be according to a law such as  $\omega^{-2}$  or  $\omega^{-4}$ ; the expected behaviour of intensity with frequency is sketched in Figure 1;



- (iii) the intensity of a pulse of radiation sent out will decay with time, possibly with the time factor  $\exp[-\nu t]$ , where  $\nu$  is the collision frequency at the point of origin.

The detection of the complete spectrum of a transient requires a device such as a spectrum analyser. However, if the minimum frequency  $\omega_0$  can be identified, an estimate of the electron density  $N$  at the place of origin may be obtained from (4). Also the collision frequency  $\nu$  in the corona is determined approximately in terms of  $N$  and the kinetic temperature  $T$  by the formula (Smerd and Westfold 3)

$$\nu = 4.2 \times 10^{-5} N T^{-3/2} \dots\dots\dots (42)$$

Hence the measured values of  $\omega_0$  and  $\nu$  enable the kinetic temperature at the place of origin to be estimated from the formula

$$T = \left( \frac{5.2 \times 10^{-7} \omega_0^2}{4\pi^2 \nu} \right)^{2/3} \dots\dots\dots (43)$$

For the other types of oscillation possible the minimum frequency and damping will depend also on the magnetic field intensity.

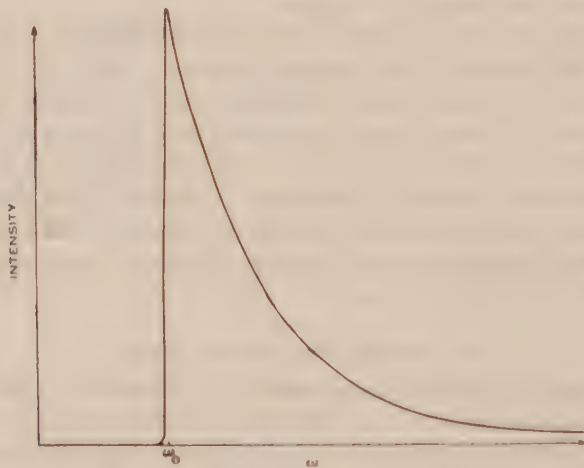


Fig. 1.—The form of spectral distribution of a transient disturbance.

We next consider the propagation of this radiation through the inhomogeneous solar atmosphere; for simplicity we consider radial propagation only. The radiation of frequency  $\omega/2\pi$  will be emitted both outwards towards the observer, and inwards towards the region of higher electron density. The wave propagated inwards will be "reflected" at the level at which the plasma frequency is  $\omega/2\pi$ , and this echo wave will arrive later than the direct wave and with less amplitude. Thus on all frequencies greater than  $\omega_0/2\pi$  we may expect to receive a "double-humped" pulse with the second hump smaller than the first; only on the frequency  $\omega_0/2\pi$  will there be a single hump. We proceed to estimate the separation in time and the difference in magnitude of the two humps. There will also be a difference in the arrival times of direct waves on



the different frequencies, the higher frequencies arriving first and the frequency  $\omega_0/2\pi$  last.

As a likely estimate of the electron density in the corona we take Allen's(4) revision of Baumbach's data

$$N = 1.42 \times 10^{14} \rho^{-6} (1 + 1.68 \rho^{-10}), \quad \rho \geq 1, \quad \dots \dots \dots (44)$$

where  $\rho$  is the distance from the centre of the sun in units of the radius  $R_1$  of the base of the corona;  $R_1 = 7.05 \times 10^5$  km. This unit is chosen for convenience in the numerical work. The refractive index of the medium for waves of frequency  $\omega/2\pi$  is then

$$\mu^2 = 1 - \frac{\omega_0^2}{\omega^2} \quad \dots \dots \dots (45)$$

with this value of  $N$ . We calculate the excess of equivalent path over distance traversed for a ray proceeding radially outwards from the point  $\rho R_1$  to infinity, namely,

$$R_1 \int_{\rho}^{\infty} \left( \frac{1}{\mu} - 1 \right) d\rho \quad \dots \dots \dots (46)$$

This is plotted in Figure 2 for frequencies of 175, 150, 100, 80, 60, and 20 Mc/s. The differences in equivalent path can be read off in units of  $R_1$ , and multiplying by 2.35 gives the time differences in seconds.

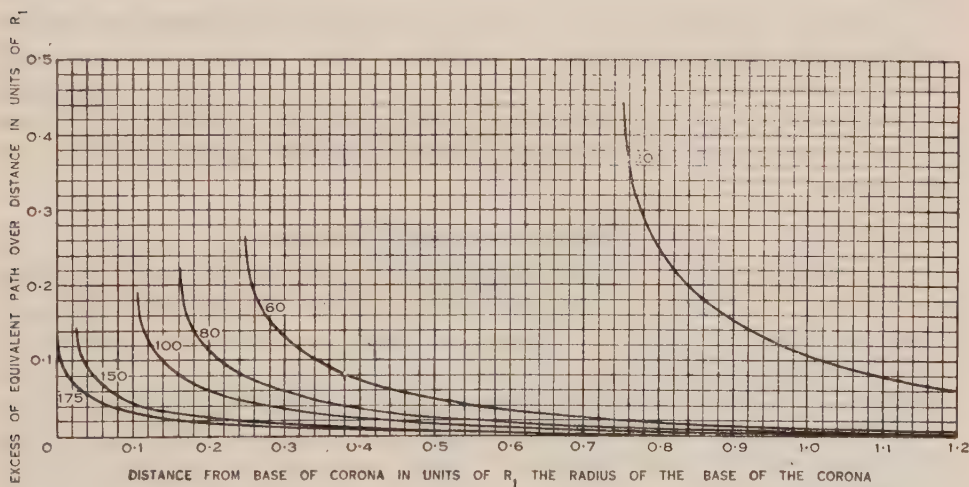


Fig. 2.—Excess of equivalent path over distance for radial propagation on various frequencies. The numbers on the curves are the frequencies in Mc/s.

As a first example, if a disturbance is produced at the 60 Mc/s. level, the direct waves on 80 and 100 Mc/s. will arrive 0.44 and 0.52 second, respectively, earlier than the wave on 60 Mc/s. The echo wave on 80 Mc/s. arrives 1.3 seconds after the direct wave, this being the delay between the two humps; on 100 Mc/s. the corresponding figure is 1.5 seconds.

If the disturbance is at the 20 Mc/s. level, the direct waves on 60, 80, and 100 Mc/s. will arrive 1.00, 1.02, and 1.04 seconds, respectively, before those on 20 Mc/s. Also the time delays between the direct and echo waves will be 3.6, 3.8, and 3.9 seconds on 60, 80, and 100 Mc/s., respectively. Because of the greater difference in path traversed the delays tend to increase with increasing frequency, but this tendency is nearly compensated by the increasing group velocity.

The absorption may be estimated by calculating

$$\int \kappa ds$$

over the trajectory of the ray, where

$$\kappa = \frac{\nu}{2c} \frac{1 - \mu^2}{\mu}, \quad \dots \dots \dots (47)$$

from the usual Lorentz theory or from (31).

For the collision frequency we use (42), taking for the kinetic temperature  $T$

$$T = 1.63 \times 10^6 \frac{\rho^{-1}(1 + 0.69\rho^{-10})}{(1 + 1.68\rho^{-10})}. \quad \dots \dots \dots (48)$$

Using these values it is found that for waves originating at the 20 Mc/s. level the ratios of the intensities of the direct and echo waves are 5 : 1, 11 : 1, and 27 : 1 on frequencies of 60, 80, and 100 Mc/s. respectively. For waves originating at the 60 Mc/s. level these ratios are 7 : 1 and 20 : 1 on 80 and 100 Mc/s.

It should be added that the values given above for radial propagation will be greatly reduced for oblique propagation. Here we have only considered radial propagation so as to indicate as simply as possible the order of magnitude of the effects. Clearly the analysis of observations on the basis of oblique incidence theory may provide considerable information about conditions in the corona.

## VIII. COMPARISON WITH EXPERIMENT

The discussion above has been confined to the study of the type of radiation received from a localized disturbance in an inhomogeneous ionized atmosphere. It has not been concerned with the way in which the disturbance is produced, and while results of the general type indicated above are always to be expected theoretically, they may be completely masked by the disturbance which produces them—for example, in the waves produced by a moving disturbance these effects may well be small compared to those produced by the motion of the disturbance itself.

If, however, any set of solar phenomena does show the general behaviour predicted in Section VII it will be reasonable to ascribe it to a localized disturbance of this type. It will be shown below that the characteristics of "bursts" of solar noise do, on the whole, satisfy this criterion.

Before doing this it is desirable to review what has already been proposed concerning the effects of steady-state plasma oscillations in the solar atmosphere. It was suggested by Martyn(5) that solar noise of intensity higher than the "quiet" level may be due to the oscillation of electron plasma.

Westfold(2) suggested that the frequencies and damping constants of bursts might be associated with those of the three types of steady-state plasma oscillations which can occur in any part of the solar atmosphere in the presence of a magnetic field. With no magnetic field the frequency is  $\omega_0/2\pi$  and the intensity decreases with the damping factor  $\exp(-\nu t)$ . If the frequency of burst radiation corresponds to one of the plasma frequencies at the level of excitation, the nearly simultaneous occurrence of a burst on several frequencies would indicate that, in effect, the cause moves with a velocity of the order of that of light. These signals should decay in intensity according to a damping factor such as  $\exp(-\nu t)$ , where  $\nu$  is the collision frequency. There will, of course, be nothing corresponding to the double-humped curves of Section VII.

An intensive study of bursts of solar noise on frequencies of 85 and 60 Mc/s. and, to a lesser extent, on 65 and 19 Mc/s. has recently been made by Ruby Payne-Scott(6), of the Division of Radiophysics, C.S.I.R.O. The following results are compared with the transient theory predictions (i), (ii), and (iii) of Section VII.

(i) Bursts show good correspondence on different frequencies with respect to both shape and time of arrival. Although the shape may vary considerably it is usually possible to discern common characteristics consistent with a common origin. There is no skipping of frequencies. Thus, if a burst appears on 85 and 19 Mc/s., it also appears on 65 and 60 Mc/s. In such circumstances the theory indicates that the level of origin is in the corona, outside the level where the plasma frequency is 19 Mc/s. In some cases a burst has appeared on 85 and 60 Mc/s. but not on 19 Mc/s., indicating a level of origin lower down in the corona.

(ii) Normally, the peak intensity of a burst on 60 Mc/s. is about twice the intensity on 85 Mc/s., indicating a variation with frequency  $\omega^{-2}$ . According to Section V this corresponds to an initial disturbance in  $\partial E/\partial t$ . The measurements on 19 Mc/s. are somewhat uncertain, because of ionospheric effects. There is no evidence of any marked increase in intensity corresponding to  $\omega=\omega_0$ ; in all but the few cases mentioned in (i) this should occur for frequencies less than 19 Mc/s.

(iii) The decay in intensity of the bursts is approximately exponential. The most common value of the damping constant on both 85 and 60 Mc/s. is  $0.6 \text{ sec.}^{-1}$ , which is of the same order of magnitude as the collision frequency in the corona. The value on 19 Mc/s. is more difficult to measure but it appears to be of the order of  $0.3 \text{ sec.}^{-1}$ . Although this value is different from that on the higher frequencies, the steady-state theory mentioned earlier would require a still smaller value.

In addition, it happens that the double-humped burst as predicted in Section VII is the usual feature. Single-humped bursts are very uncommon. On the 85 and 60 Mc/s. records that have been measured, the larger hump arrives first, followed about 2 to 6 seconds later by the smaller hump. The peak intensities are in the ratio of about 4 : 1. These results are in reasonable agreement with the values calculated in Section VII for radiation originating at the 20 Mc/s. level.

Normally, the arrival of bursts on different frequencies is in the temporal order predicted in Section VII, but the observed delays of about 0.7 and 9 seconds between 85 and 60, and 60 and 19 Mc/s. are too great.

On rare occasions bursts do not have the characteristics that have been designated as "normal". Then, over a period of several hours, bursts on 60 Mc/s. arrive *before* bursts on 85 Mc/s. Such bursts are evidently outside the scope of the theory presented in this paper.

#### IX. ACKNOWLEDGMENT

The work described in this paper was carried out as part of the research programme of the Division of Radiophysics, C.S.I.R.O.

#### X. REFERENCES

- (1) CARSLAW, H. S., and JAEGER, J. C.—"Operational Methods in Applied Mathematics" 2nd Ed. (Oxford Univ. Press: Oxford, 1948.)
- (2) WESTFOLD, K. C.—*Aust. J. Sci. Res. A* **2**: 169-83 (1949).
- (3) SMERD, S. F., and WESTFOLD, K. C.—*Phil. Mag.* **40**: 831-48 (1949).
- (4) ALLEN, C. W.—*Mon. Not. Roy. Astr. Soc.* **107**: 426-32 (1947).
- (5) MARTYN, D. F.—*Nature* **159**: 26-7 (1947).
- (6) PAYNE-SCOTT, RUBY.—*Aust. J. Sci. Res. A* **2**: 214-27 (1949).



# THE DISTRIBUTION OF LIGHT FROM OPTICAL SYSTEMS

By W. H. STEEL\*

(Plates 1-3)

[*Manuscript received March 30, 1949*]

## *Summary*

A geometrical method is developed for calculating the distribution of intensity with angle of the light leaving an optical system, when the angle at which a ray from a point on the optical axis leaves the system is known as a function of the aperture. The case of a point source on the axis of the system is treated exactly, and an approximation is given for that of a small finite source; the method is applicable to systems with spherical aberration. The distribution of illumination across any plane normal to the axis is treated by analogous methods. The results are compared with measurements of the light distribution from an optical system possessing considerable spherical aberration.

## I. INTRODUCTION

In its practical applications, the problem of the light distribution produced by an optical system may be arbitrarily divided into two general cases: the distribution near an image, and that at points remote from an image plane, as, for example, in light projection systems. The complete solution in either case requires the use of wave optics, and the distribution of light near the image has been studied by several authors (e.g. 1-5) by means of numerical or mechanical integrations in terms of the phase differences across the emergent wave front. However, as considerable information can be obtained from the "smoothed" light distribution given by geometrical optics, in which interference and the consequent "fine structure" of diffraction are ignored, such a treatment is developed in this paper. For projection equipment used well out of focus, wave optics methods would be very complicated, and in practice the diffraction effects would not be noticeable across the wide distribution involved. In such instances the geometrical approach gives all the information required.

Several geometrical treatments have been developed for various aspects of the problem. Hawkins and Linfoot(6) and Herzberger(7) have studied well-corrected systems using spot diagrams, the spots representing the intersection of the image plane by rays evenly spaced in the aperture of the system. Such a treatment provides no general solution but can be used to give a pictorial representation of the light distribution in the image of a point source either on or off the axis in the presence of any type of aberration. Blondel and Lavanchy(8, 9) have considered the various projection systems that are used in lighthouses;

\* Division of Physics, C.S.I.R.O.

they have dealt with a finite source at or near the focus of a system substantially free from aberration. Their method has been analysed and extended by Dunoyer(10).

In this paper, a mathematical analysis is given for the light distribution from a source on the optical axis; it is applicable to systems with spherical aberration. The calculation of this distribution is based on the results of a set of ray traces for an axial pencil\*; alternatively, approximate algebraic expressions based on third order aberration theory are obtained for the light distribution, and used to study its general form. The analysis for a point source on the axis is treated first, but as the application of geometrical optics to a point source gives infinite values in those directions where maxima occur, it is necessary to extend the discussion to the case of a small but finite source, thereby obtaining a complete solution.

A publication has recently come to hand containing an article by van Heel(11) in which the case of a point source has been treated. The methods used and the results obtained for such a source differ only in detail from those of the present paper.

The notation used is, as far as possible, that of Conrady(12), where the angles  $U$  or  $U'$  between an object or image ray and the optical axis are of opposite sign to the cartesian coordinate convention.

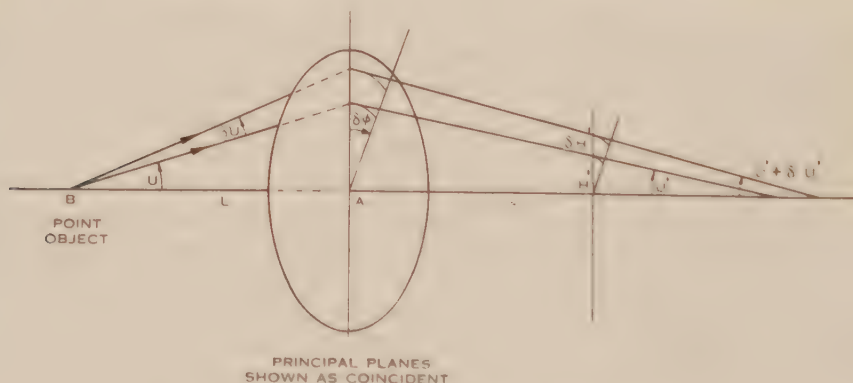


Fig. 1.—Light distribution from a point source.

## II. LIGHT DISTRIBUTION FROM A POINT OBJECT

Referring to Figure 1, the object  $B$  is a point source of finite intensity  $I$  on the optical axis of the system at a distance  $L$  from the object principal plane.

Consider light leaving the source at angles between  $U$  and  $U + \delta U$  to the optical axis in tangential planes at azimuth between  $\varphi = 0$  and  $\varphi = \delta\varphi$ . The solid angle is  $\delta U \delta\varphi \sin U$ , and, if the source has a uniform intensity in all directions, the element of flux entering this solid angle is  $\delta F = I \delta U \delta\varphi \sin U$ .

\* An "axial pencil" is used as in Conrady (12, Chapter VI) to denote a pencil of rays from a point on the optical axis of the system.

(a) *Intensity Due to the System in a Given Direction*

Corresponding to the above, flux  $t\delta F$ , where  $t$  is the transmission of the system, leaves at angles between  $U'$  and  $U' + \delta U'$  to the optical axis, where  $U'$  is a function  $U'(U)$  of  $U$ , and  $U' + \delta U' = U'(U + \delta U)$ .

This solid angle is

$$\delta\omega' = \delta U' \delta\varphi \sin U',$$

and the intensity of the system in the direction  $U'$  is given by

$$I' = t \frac{dF}{d\omega'} ,$$

$$= tI \frac{\sin U}{\sin U'} \cdot \frac{dU}{dU'} ,$$

provided  $\delta U/\delta U'$  tends to a finite limit  $dU/dU'$  as  $\delta U \rightarrow 0$ . This derivative is actually  $\left(\frac{\partial U}{\partial U'}\right)_{L \text{ constant}}$ ; it is infinite at a focus, and such cases are treated later in this paper.

To obtain a property of the system independent of the intensity of the source, we define, following Van Lear(13),  $i$  the specific intensity as  $I'/E_0$ , where  $E_0$  is the illumination at the centre of the principal plane of the system.

Then

$$E_0 = I/L^2,$$

$$i = tL^2 \frac{\sin U}{\sin U'} \cdot \frac{1}{dU'/dU} \dots\dots\dots (1)$$

We can consider the intensity in terms of the aperture  $Y$  of the ray, where

$$Y = L \tan U ;$$

then

$$i = t \frac{Y}{\sin U'} \cdot \frac{\cos^3 U}{dU'/dY} \dots\dots\dots (2)$$

(b) *Illumination on a Plane at a Finite Distance from the System*

Take any plane normal to the optical axis at a finite distance from the system. If the rays leaving the source at angles  $U$  and  $U + \delta U$  cut this plane at distances  $H'$  and  $H' + \delta H'$  respectively from the optical axis, then the flux  $t\delta F$  falls on an area  $\delta A' = H' \delta H' \delta\varphi$ .

The illumination  $E'$  at a distance  $H'$  from the axis is thus

$$E' = t \frac{dF}{dA'} ,$$

$$= tI \frac{\sin U}{H'} \cdot \frac{dU}{dH'} ,$$

provided  $dU/dH'$  is finite. If we write  $\epsilon = E'/E_0$ , then

$$\epsilon = tL^2 \frac{\sin U}{H'} \cdot \frac{1}{dH'/dU} ,$$

or, in terms of aperture,

$$\epsilon = t \frac{Y}{H'} \cdot \frac{\cos^3 U}{dH'/dY} \dots\dots\dots (3)$$

where  $dH'/dY$  is the derivative of the height  $H'$  at which an axial ray cuts the given plane with respect to the aperture of this ray.

This expression differs from that of van Heel, who obtained the factor  $\cos^3 U'$  in place of  $\cos^3 U$ .

### III. DISTRIBUTION OF LIGHT FROM A FINITE OBJECT

#### (a) *The Intensity in a Given Direction*

Expressions (1) and (2) do not give finite values when either  $U'$  or  $dU'/dY$  vanishes. However, these cases can be treated when account is taken of the finite size of the source. The treatment is similar to the method used by Blondel and Lavanchy, where the area of the principal plane of the system that sends light from the source at an angle  $U'$  to the axis is calculated.

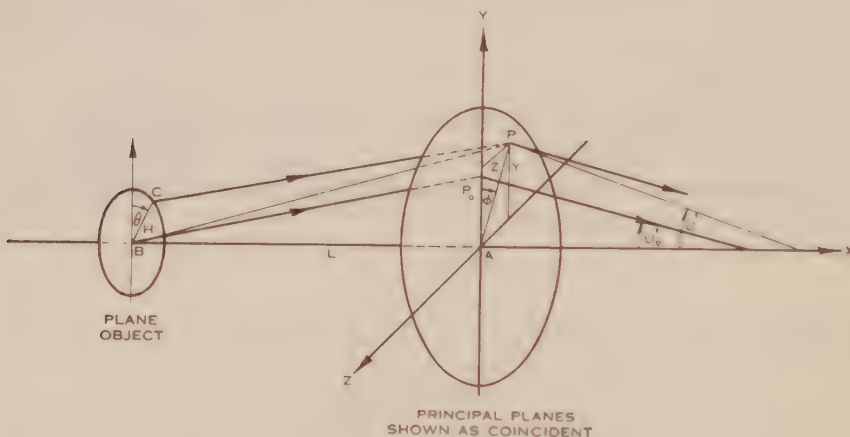


Fig. 2.—Intensity distribution from a finite source.

In Figure 2, a circular object of radius  $H$  and brightness  $B$  is centred on and normal to the optical axis at a distance  $L$  from the principal plane of an optical system. We require to find the intensity due to the system at an angle  $U'_0$  to the optical axis or the illumination on a given plane at a distance  $H'_0$  from the axis, both in an azimuth plane  $\psi=0$ .

Take the optical axis as  $X$  coordinate, with the  $Y$  and  $Z$  coordinates at right angles so that  $OY$  lies in the plane  $\psi=0$ . Then if  $\lambda, \mu, \nu$  are the direction cosines of the incident ray, and  $\lambda', \mu', \nu'$  those for the ray leaving the system (14),

$$\frac{\mu - \mu'}{Y} = \frac{\nu - \nu'}{Z} = K,$$

where  $(O, Y, Z)$  is the point of intersection of the ray with the object principal plane, assuming the principal surfaces to be plane, and  $K$  is the power of the lens for the ray, not in general a constant but a function of  $\sqrt{Y^2 + Z^2}$  and of the direction of the ray.

In general, for each point on the object, there is a corresponding point of the principal plane from which the emergent ray is in the required direction



$U'_0$ . Corresponding to the centre of the object we have the point  $P_0$ ,  $(O, Y_0, O)$ , and  
 and  $\mu = -\sin U_0$ ,  $\mu' = -\sin U'_0$ ,  
 so that  $\sin U'_0 - \sin U_0 = K_0 Y_0$

If  $C$  is a point in the edge of the object at azimuth  $\theta$ , and  $P(O, Y, Z)$  is the corresponding point of the principal plane at azimuth  $\varphi = \text{artan } Z/Y$ , then

$$\mu = \frac{Y - H \cos \theta}{D}, \quad \nu = \frac{Z - H \sin \theta}{D},$$

$$\mu' = -\sin U'_0, \quad \nu' = 0,$$

where  $D$  is the distance  $CP$ , so that

$$\left. \begin{aligned} \sin U'_0 + \frac{Y}{D} - \frac{H}{D} \cos \theta &= KY \\ \frac{Z}{D} - \frac{H}{D} \sin \theta &= KZ \end{aligned} \right\} \dots\dots\dots (4)$$

If the function  $K$  is known, we can solve for  $Y$  and  $Z$ . Then as  $C$  moves around the edge of the source from  $\theta=0$  to  $\theta=2\pi$ ,  $P$  traces out a closed curve at the principal plane of the system. If this curve has an area  $\delta A$ , the projected area normal to the direction  $U'_0$  is  $\delta A \cos U'_0$ . The brightness of any area normal to a ray through the system from a point on the source is  $tB_u$  (15) if both the source and this area are in media of the same refractive index,  $t$  being the transmission of the optical system between the source and this area, and  $B_u$  the brightness of the source in the direction in which the ray leaves it. For a source of uniform brightness,  $B_u = B$ , a constant for all angles, and the brightness of this area  $\delta A \cos U'_0$  is  $tB$ . Hence the intensity due to the system in the direction  $U'_0$  is  $tB\delta A \cos U'_0$ .

Usually the function  $K$  is known in terms of the variation of  $U'$  with  $Y$  for an axial pencil. We therefore compare these equations (4) with those for a tangential ray through the same point  $P$ , but from the centre of the object, namely,

$$\left. \begin{aligned} \sin U' \cos \varphi - \sin U \cos \varphi &= KY, \\ \sin U' \sin \varphi - \sin U \sin \varphi &= KZ. \end{aligned} \right\} \dots\dots\dots (5)$$

If the object has a small angular size  $e = H/L$ , the difference between the direction of this incident ray and the previous one is small, so the same value of  $K$  can be assumed for both equations (4) and (5).

Also  $D^2 = (L^2 + Y^2 + Z^2)(1 + 2e \cos U \sin U \cos \overline{\theta - \varphi} + e^2 \cos^2 U)$ , and using this to expand equations (4) in terms up to the first order in  $e$ , and combining them with equations (5), we obtain

$$\sin U'_0 - e \cos U \cos \theta + e \sin^2 U \cos U \cos \varphi \cos \overline{\theta - \varphi} = \sin U' \cos \varphi, \quad \dots\dots\dots (6)$$

$$-e \cos U \sin \theta + e \sin^2 U \cos U \sin \varphi \cos \overline{\theta - \varphi} = \sin U' \sin \varphi, \quad \dots\dots\dots (7)$$

$$\text{giving} \quad \sin U'_0 \sin \varphi + e \cos U \sin \overline{\theta - \varphi} = 0. \quad \dots\dots\dots (8)$$

These can be solved simply for special values of  $\theta$ .

(i) When  $\theta=0$ ,  $P$  lies in the plane  $\varphi=0$ , i.e.  $Z=0$ , and then equation (6) gives

$$\sin U'_0 - e \cos^3 U = \sin U'.$$

But

$$\sin U' = \sin U'_0 + (U' - U'_0) \cos U'_0 + \dots$$

so, to the first order in  $e$ ,  $U' - U'_0 = -e \frac{\cos^3 U}{\cos U'_0} \dots \dots \dots (9)$

But, as  $U'$  is a function of  $Y$ , and  $U'_0$  its value for  $Y = Y_0$ ,

$$U' - U'_0 = (Y - Y_0) \left( \frac{dU'}{dY} \right)_{Y_0} + \frac{(Y - Y_0)^2}{2!} \left( \frac{d^2 U'}{dY^2} \right)_{Y_0} + \frac{(Y - Y_0)^3}{3!} \left( \frac{d^3 U'}{dY^3} \right)_{Y_0} + \dots, \dots \dots (10)$$

where  $\left( \frac{d^n U'}{dY^n} \right)_{Y_0}$  is the value of the derivative at  $Y = Y_0$ .

If  $\left( \frac{dU'}{dY} \right)_{Y_0}$  is not zero, we have for the point  $P$

$$Y \doteq Y_0 - \frac{e \cos^3 U}{\cos U'_0 \left( \frac{dU'}{dY} \right)_{Y_0}}.$$

$$(ii) \text{ When } \theta = \pi, Y \doteq Y_0 + \frac{e \cos^3 U}{\cos U'_0 \left( \frac{dU'}{dY} \right)_{Y_0}}.$$

(iii) When  $\theta = \pi/2$ , to the first order in  $e$ ,  $Y \doteq Y_0$

and equation (8) gives  $\tan \varphi = -\frac{e \cos U}{\sin U'_0}$ ,

so  $Z \doteq -Y_0 \frac{e \cos U}{\sin U'_0}$ .

$$(iv) \text{ When } \theta = 3\pi/2, Z \doteq Y_0 \frac{e \cos U}{\sin U'_0}.$$

The value of  $\cos U$  differs for each case, but for  $e$  small we can replace it by  $\cos U_0$ ,  $U_0$  being the angle with the axis for a ray from the centre of the object to  $P_0$ ,  $\tan U_0 = Y_0/L$ .

Assuming that  $\delta A$  is approximately of the same area as an ellipse of semi-axes,

$$Y_0 \frac{e \cos U_0}{\sin U'_0} \text{ and } \frac{e \cos^3 U_0}{\cos U'_0 \left( \frac{dU'}{dY} \right)_{Y_0}},$$

and since the projected area is  $\delta A \cos U'_0$ ,

$$I' = tB \frac{e^2 Y_0}{\sin U'_0} \cdot \frac{\cos^4 U_0}{\left( \frac{dU'}{dY} \right)_{Y_0}}.$$

For  $e$  small, the illumination at the centre of the principal plane of the system is  $E_0 = \pi B e^2$  and the specific intensity is given by

$$i = t \frac{Y_0}{\sin U'_0} \cdot \frac{\cos^4 U_0}{\left( \frac{dU'}{dY} \right)_{Y_0}} \dots \dots \dots (11)$$

The intensity due to a small plane source thus approximates to that for a point source; the fourth power of  $\cos U_0$  now appears because the previous treatment started from a point source of uniform intensity, while a plane source of uniform brightness becomes in the limit a point of intensity  $I=I_0 \cos U$ .

If the source is not small, the area  $\delta A$  may be very different from an ellipse, and the actual shape would have to be considered in detail by methods such as those employed by Dunoyer(10).

(b) *The Illumination at a Given Point on a Given Plane*

The illumination at any point can be obtained by a similar method. If the point lies on a plane at a distance  $s$  from the principal plane of the system and is at a distance  $H'_0$  from the optical axis, the direction cosines of the ray leaving the system now must be

$$\mu' = \frac{H'_0 - Y}{D'}, \quad \nu' = -\frac{Z}{D'},$$

where  $D'$  is the distance from the point  $P$  to the point at which the illumination is required.

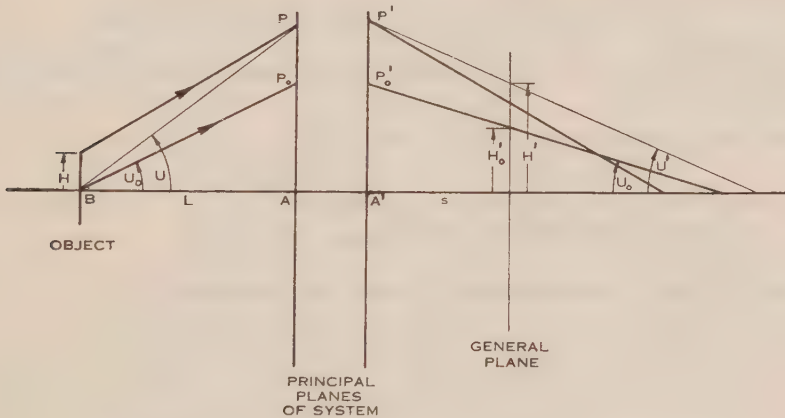


Fig. 3.—Illumination distribution from a finite source—  
path of rays in tangential plane.

We consider again a tangential ray from the centre of the source through the point  $P$ . This ray cuts the plane at a distance  $H'$  from the axis in an azimuth plane  $\psi = \varphi$ .

On comparing the equations for these two rays, we can obtain the values :

(i) When  $\theta = \varphi = 0$ , as shown in Figure 3,

$$Y - Y_0 = \frac{se \cos^3 U}{\cos^3 U'} \cdot \frac{1}{\left(\frac{dH'}{dY}\right)_{Y_0}};$$

(ii) When  $\theta = \frac{\pi}{2}$ ,  $Z = \frac{Y_0 se \cos U}{H'_0 \cos U'}$ .

The area  $\delta A$  of the principal plane that sends light to the point considered is approximately that of an ellipse of these semi-axes, and, if this area is small, the illumination  $E'$  at the point is

$$\frac{tB\delta A \cos^4 U'_0}{s^2}.$$

Again the values of  $\cos U$  and  $\cos U'$  differ for each point on this area, but we can replace them by  $\cos U_0$  and  $\cos U'_0$  if  $\delta A$  is small.

$$\text{Then } E' = \frac{tB \cos^4 U'_0}{s^2} \cdot \frac{Y_0 s e \cos U_0}{\cos U'_0} \cdot \frac{s e \cos^3 U_0}{\cos^3 U'_0 \left( \frac{dH'}{dY} \right)_{Y_0}},$$

$$\text{and } \varepsilon = \frac{E'}{E_0} = \frac{tY_0}{H'_0} \cdot \frac{\cos^4 U_0}{\left( \frac{dH'}{dY} \right)_{Y_0}} \dots\dots\dots (12)$$

#### IV. SPECIAL CASES

We now consider the cases where the above discussion leads to infinite values.

##### (a) *The Intensity or the Illumination on a Caustic*

When  $\left( \frac{dH'}{dY} \right)_{Y_0} = 0$ ,  $H'$  has a stationary value which we shall denote by  $R$ , and the point is on a caustic; when  $\left( \frac{dU'}{dY} \right)_{Y_0} = 0$ ,  $U'$  has a stationary value,  $a$ , and this is the direction of a caustic at infinity. In these cases we must use a second derivative to approximate to  $Y - Y_0$ .

When  $\left( \frac{dU'}{dY} \right)_{Y_0} = 0$ , equations (9) and (10) give us

$$Y - Y_a = \pm \sqrt{\frac{-2 e \cos^3 U_0}{\cos a \left( \frac{d^2 U'}{dY^2} \right)_{Y_a}}},$$

where  $Y_a$  is the value of  $Y_0$  when  $U'_0$  has a stationary value  $a$ .

If  $a$  is a maximum value of  $U'$ ,  $\left( \frac{d^2 U'}{dY^2} \right)_{Y_a}$  is negative, so there is no point  $P$  corresponding to  $\theta = \pi$ , and, in fact, one-half only of the source sends light in the required direction. The area  $\delta A$  is still approximately that of an ellipse as each point on this half of the source sends light through two points at the principal plane, corresponding to the two opposite values of  $Y - Y_a$ .

The semi-axes are  $\frac{eY_a \cos U_a}{\sin a}$  as before

and

$$\sqrt{\frac{-2e \cos^3 U_a}{\cos a \left( \frac{d^2 U'}{dY^2} \right)_{Y_a}}}.$$



and 
$$i = \frac{tY_a \cos^2 U_a}{\tan a} \sqrt{\frac{-2 \cos U_a}{e \cos a \left(\frac{d^2 U'}{dY^2}\right) Y_a}} \dots\dots\dots (13)$$

It will be seen later, when the shape of the function  $i(U')$  is considered in terms of primary aberration, that there is a maximum value of  $i$  near but not at  $U'=a$ .

The angle closest to  $a$  into which the whole source sends light is approximately  $a - e \frac{\cos^3 U_a}{\cos a}$ . If  $e$  is small compared with  $a$  we can use the same values of the derivatives at this angle as at  $a$ . Then for the centre of the source.

$$Y_0 = Y_a \sqrt{\frac{-2e \cos^3 U_a}{\cos a \left(\frac{d^2 U'}{dY^2}\right) Y_a}},$$

and for  $\theta=0$ , 
$$Y - Y_0 = \pm \sqrt{\frac{-2e \cos^3 U_a}{\cos a \left(\frac{d^2 U'}{dY^2}\right) Y_a}}.$$

The area is now approximately that of two touching ellipses of semi-axes

$$\frac{e \cos U_0}{\sin\left(a - e \frac{\cos^3 U_a}{\cos a}\right)} \left\{ Y_a \pm \sqrt{\frac{-2e \cos^3 U_a}{\cos a \left(\frac{d^2 U'}{dY^2}\right) Y_a}} \right\} \text{ and } \sqrt{\frac{-2e \cos^3 U_a}{\cos a \left(\frac{d^2 U'}{dY^2}\right) Y_a}},$$

and 
$$i = \frac{2tY_a \cos^2 U_a}{\tan\left(a - e \frac{\cos^3 U_a}{\cos a}\right)} \cdot \sqrt{\frac{-2 \cos U_a}{e \cos a \left(\frac{d^2 U'}{dY^2}\right) Y_a}}, \dots\dots\dots (14)$$

which, if  $e$  is small, is approximately twice the value for  $U'=a$ .

However, as the value of the first derivative cannot in general be ignored, the maximum value of  $i$  occurs for  $U'$  between  $a$  and  $a - e \frac{\cos^3 U_a}{\cos a}$  and lies between the values given by equations (13) and (14).

The corresponding value for the illumination at a stationary value of  $R$  of  $H'$  is

$$\varepsilon = \frac{tY_R \cos U'_R \cos^2 U_R}{R} \sqrt{\frac{2 \cos U'_R \cos U_R}{se \left(\frac{d^2 H'}{dY^2}\right) Y_R}} \dots\dots (15)$$

The nearest value of  $H'$  to  $R$  to which the whole source sends light is

$$H' = R - se \frac{\cos^3 U_R}{\cos^3 U'_R},$$

and there

$$\varepsilon = \frac{2tY_R \cos U'_R \cos^2 U_R}{R - se \frac{\cos^3 U_R}{\cos^3 U'_R}} \cdot \sqrt{\frac{2 \cos U'_R \cos U_R}{se \left(\frac{d^2 H'}{dY^2}\right) Y_R}}, \dots (16)$$

the maximum value of  $\varepsilon$  occurring between the values given by (15) and (16) for  $H'$  between  $R$  and  $R - se \frac{\cos^3 U_R}{\cos^3 U'_R}$ .

(b) *The Intensity or the Illumination along the Axis when  $U' = 0$  or  $H' = 0$ .*

The ray leaves the system at angle  $U' = 0$  for  $Y_0 = 0$  and also for some other values of  $Y_0^2$  that may give real solutions  $Y_0 = \pm b$ .

Near  $Y_0 = 0$ , the value of  $Y$  for the edge of the source is  $\left(\frac{e}{dY}\right)_0$  for all values

of  $\theta$ . Then  $\delta A$  is the area of a circle of this radius and

$$i = \frac{t}{\left(\frac{dU'}{dY}\right)_0^2} \dots \dots \dots (17)$$

If  $b$  is real, the corresponding values of  $Y$  are

$$Y = b \pm \frac{e \cos^3 U_0}{\left(\frac{dU'}{dY}\right)_b},$$

for all values of  $\theta$ , so  $\delta A$  is the area of an annulus of mean radius  $b$  and width

$$\frac{2e \cos^3 U_0}{\left(\frac{dU'}{dY}\right)_b},$$

whence

$$i = \frac{4tb \cos^3 U_0}{e \left(\frac{dU'}{dY}\right)_b} \dots \dots \dots (18)$$

so that the total intensity along the axis is the sum of the expressions (17) and (18).

The illumination for a point on the axis is similarly

$$i = \frac{t}{\left(\frac{dH'}{dY}\right)_0^2} + \frac{4tb \cos U'_0 \cos^3 U_0}{se \left(\frac{dH'}{dY}\right)_b} \dots \dots \dots (19)$$

(c) *The Intensity or the Illumination at Paraxial Focus*

If  $U' = 0$  and  $b = 0$ , we have three roots of  $U'(Y) = 0$  at  $Y = 0$ , and  $\left(\frac{dU'}{dY}\right)_0 = \left(\frac{d^2 U'}{dY^2}\right)_0 = 0$ . This is the case of a paraxial image at infinity.

The area is again a circle, radius

$$Y = \left[ \frac{-6e}{\left(\frac{d^3 U'}{dY^3}\right)_0} \right]^{1/3},$$

and

$$i = t \left[ \frac{-6}{e^2 \left(\frac{d^3 U'}{dY^3}\right)_0} \right]^{2/3} \dots \dots \dots (20)$$

The illumination at a paraxial focus on the axis is similarly

$$\varepsilon = t \left[ \frac{6}{s^2 c^2 \left( \frac{d^3 H'}{dY^3} \right)_0} \right]^{2/3}, \quad \dots\dots\dots (21)$$

where  $H'$  is now the tangential spherical aberration  $TA'$ .

## V. APPLICATION TO CORRECTED SYSTEMS

The previous theory is not suitable for the treatment of the illumination at the focus of a corrected system, or the intensity at an image at infinity, where the flashed area  $\delta A$  is large and limited, at least in part, by the edge of the system. These cases are treated by Blondel and Lavanchy(8, 9), and by Dunoyer(10). However, the illumination or the intensity can be obtained away from the focus.

(i) For a perfect system where

$$L \tan U = L' \tan U' = Y,$$

the image distance  $L'$  being constant, the intensity at an angle  $U'$  is

$$i = t L'^2 \frac{\cos^4 U}{\cos^3 U'}.$$

The illumination on a plane at a distance  $s$  along the optical axis is

$$\varepsilon = t \left( \frac{L'}{L' - s} \right)^2 \cos^4 U.$$

(ii) For a system obeying the sine condition,

$$\sin U' = m \sin U, \text{ where } m \text{ is constant,}$$

$$i = \frac{t L^2}{m^2} \cos U',$$

the distribution following a cosine law, as does the source.

## VI. PRIMARY ABERRATION TREATMENT OF INTENSITY

Analytic expressions can be obtained for the intensity using an approximate algebraic expression for the spherical aberration of the system, which gives  $U'$  as a function of  $Y$ .

Such expressions are only accurate for small angles when  $\cos U \doteq 1$ ,  $\sin U \doteq U$ , etc. We shall use  $u$  and  $y$ , but these do not denote a true paraxial treatment, as the primary aberration is included.

If the spherical aberration is not corrected, the angular aberration which can be expanded as

$$AA' = c_1 y^3 + c_2 y^5 + \dots$$

is predominantly that of the leading term  $c_1 y^3$ . Higher power terms alter the form of the function slightly, but do not introduce extra roots of  $U'(y) = 0$ , where

$$U'(y) = \frac{y}{l'} + AA',$$

unless the spherical aberration has been substantially corrected.

If we consider only primary aberration,

$$u'(y) = \frac{y}{l'} + cy^3,$$

where  $y/l'$  is the paraxial value of  $U'$ ,  $l'$  being the distance of the paraxial image of the source from the image principal plane of the system.

Assuming perfect transmission,  $t=1$ , equation (6) reduces to

$$i' = \frac{y}{u'} \cdot \frac{1}{du'/dy} \dots\dots\dots (22)$$

Eliminating  $y$ , we obtain an implicit expression for  $i$  in terms of  $u'$

$$27l'^3c^2i^3u'^4 + 4ci^3u'^2 + l'i - l'^3 = 0 \dots\dots\dots (23)$$

and the shape of the function  $i(u')$  depends on the values of  $l'$  and  $c$ . This equation applies to any system when the spherical aberration is predominantly that of the first order.

#### (a) No Compensation of Aberration

If  $l'$  and  $c$  are the same sign,  $u'$  increases or decreases steadily with  $y$ .  $i$  is a maximum for  $u'=0$  where it equals  $l'^2$ , dropping off to zero at  $u' = \pm \infty$ . The shape is shown in curve A, Figure 4.

#### (b) Paraxial Focus at Infinity

If  $l' = \infty$ , the paraxial image is at infinity and

$$i = \frac{1}{3}c^{-2/3}u'^{-4/3} \text{ (curve B, Fig. 4)}$$

$i$  increases without limit as  $u'$  approaches 0. When  $u' = 0$  for a source of finite size and brightness, by equation (20),

$$i = c^{-2/3}c^{-4/3}.$$

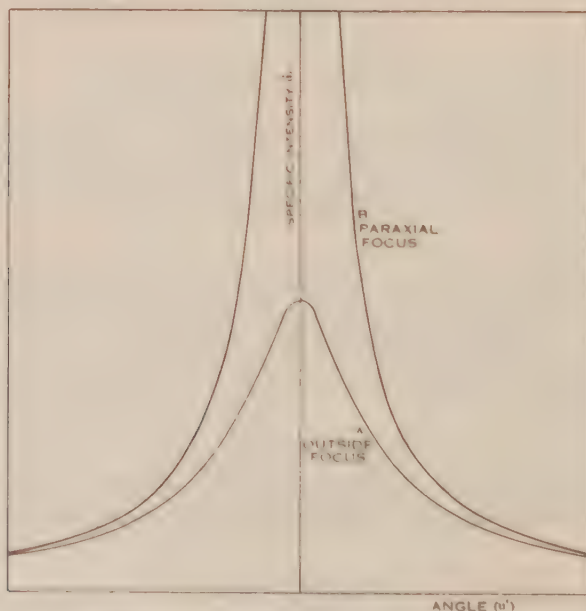


Fig. 4.—Light distribution at and outside paraxial focus.



(c) *Aberration Compensated by Change of Focus*

If  $l'$  and  $c$  are of opposite sign the spherical aberration is compensated at some aperture by a change of focus so that  $u'=0$  for a zone  $y=\pm b$  as well as for  $y=0$ .

Rather than use  $l'$  and  $c$  as calculated by algebraic third order treatment which gives an approximation near  $y=0$  we can give them values that will fit the results of ray traces at certain finite values of  $y$ . Choosing the two conditions :

$$(i) \quad u'=0, \text{ when } y=b,$$

$$(ii) \quad du'/dy=0, \text{ when } u'=a,$$

where  $a, b$  are given by ray traces,

$$u' = \frac{3\sqrt{3}}{2} a \left( \frac{y^3}{b^3} - \frac{y}{b} \right),$$

and equation (23) becomes

$$729a^2i^3u'^4 - 729a^4i^3u'^2 + 27a^2b^4i - 4b^6 = 0.$$

$i$  is now a triple valued function of  $u'$  and the intensity in any direction  $u'$  is the sum of the values of  $|i|$  for that  $u'$ . In Figure 5, the function  $i(u')$  is plotted with the dotted line; the full curve is  $\Sigma |i|$ .

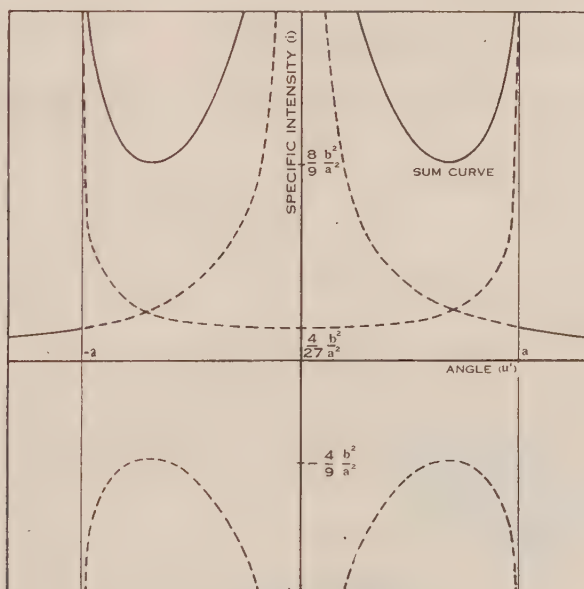


Fig. 5.—Light distribution inside paraxial focus.

The intensity shows peaks at  $u'=0$ , and at  $|u'|=a$ , i.e. a bright centre and a surrounding bright ring. Beyond  $|u'|=a$  the intensity is small and decreases rapidly as  $|u'|$  increases. When  $u'=a$ ,  $y=-\frac{b}{\sqrt{3}}$ ,  $-\frac{b}{\sqrt{3}}$ , or  $+\frac{2b}{\sqrt{3}}$ , so that the aperture beyond  $|y|=\frac{2b}{\sqrt{3}}$  contributes no light to the central concentration within  $|u'| \leq a$ .

Between the centre and the "bright ring" there is a minimum intensity when  $u' = \frac{a}{\sqrt{2}}$ , i.e. when  $y = -\sqrt{\frac{2}{3}} a$ , or  $\frac{a}{\sqrt{2}} \left( \frac{1}{\sqrt{3}} \pm 1 \right)$ .

Then 
$$i_{\min} = \frac{8}{9} \frac{b^2}{a^2}.$$

We can also find the average intensity within  $|u'| < a$ . The area sending light within this angle is  $\frac{4}{3}\pi b^2$  and the solid angle into which this light goes is  $\pi a^2$ , so that

$$i_{av} = \frac{4}{3} \frac{b^2}{a^2},$$

$$= \frac{3}{2} i_{\min}.$$

(d) *Primary Aberration Treatment of Special Cases*

The curve of Figure 5 gives  $i$  infinite at  $u' = 0, \pm a$ . We can find  $i$  at these angles using the special treatment above.

(i) At  $u' = a, y = -\frac{b}{\sqrt{3}}, -\frac{b}{\sqrt{3}}, +\frac{2b}{\sqrt{3}}.$

Applying equation (13) for the roots  $y = -\frac{b}{\sqrt{3}}$ , and equation (22) for the single root  $\frac{2b}{\sqrt{3}},$

$$i = \frac{b^2}{a^2} \left( \frac{4}{27} + \frac{2}{3\sqrt{3}} \sqrt{\frac{a}{e}} \right).$$

(ii) The maximum value for  $i$  lies between this and the value given by equation (14), namely,

$$i = \frac{b^2}{a^2} \left( \frac{4}{27} + \frac{2}{3\sqrt{3}} \sqrt{\frac{a}{e}} \right).$$

It is less than this latter value, as the first derivative does not remain zero. This maximum occurs between  $u' = a$  and  $u' = a - e$ .

(iii) At  $u' = 0, y = 0, \pm b.$

Combining equations (17) and (18),

$$i = \frac{b^2}{a^2} \left( \frac{4}{27} + \frac{4}{3\sqrt{3}} \sqrt{\frac{a}{e}} \right).$$

VII. PRIMARY ABERRATION TREATMENT OF ILLUMINATION

In terms of primary aberration,  $u' = \frac{y}{f} + cy^2.$

so that

$$h' = y - su' \text{ for small angles,}$$

$$= y \left( 1 - \frac{s}{f} - cxy^2 \right),$$

and

$$\epsilon = \frac{y}{h'} \cdot \frac{1}{dh'/dy}.$$

The distribution takes the same form as the intensity, the shape depending on

whether  $\frac{1-\frac{s}{l'}}{cs}$  is positive, 0, or negative. If this expression for  $y$  is fitted to the conditions :

$$(i) \quad h'=0, \text{ when } y=b,$$

$$(ii) \quad dh'/dy=0, \text{ when } h'=r,$$

where  $r, b$  are given by ray traces,

$$h' = \frac{3\sqrt{3}}{2} r \left( \frac{y^3}{b^3} - \frac{y}{b} \right).$$

We can list the following values of  $\epsilon$ , again assuming the transmission to be perfect :

$$(i) \text{ At } h'=r, \quad \epsilon = \frac{b^2}{r^2} \left( \frac{4}{27} + \frac{\sqrt{2}}{3\sqrt{3}} \sqrt{\frac{r}{se}} \right).$$

$$(ii) \text{ At } h'=r-se, \quad \epsilon = \frac{b^2}{r^2} \left( \frac{4}{27} + \frac{2\sqrt{2}}{3\sqrt{3}} \sqrt{\frac{r}{se}} \right).$$

$$(iii) \text{ At } h'=0, \quad \epsilon = \frac{b^2}{r^2} \left( \frac{4}{27} + \frac{4}{3\sqrt{3}} \cdot \frac{r}{se} \right).$$

$$(iv) \text{ Minimum value } \epsilon_{min} = \frac{8}{9} \frac{b^2}{r^2}.$$

$$(v) \text{ Average value, for } |h'| \leq r, \quad \epsilon_{av} = \frac{3}{2} \epsilon_{min}.$$

### VIII. COMPARISON WITH EXPERIMENT

Measurements were made on an uncorrected optical system to compare the results with those predicted by this theory. The system used was a "reflector button" as employed in traffic warning signals. It consisted of a thick double convex lens with a back surface silvered and at the principal focus of the front surface. Such a system possesses spherical aberration and for parallel incident light the reflected light is distributed around the incident direction. The system is shown in Figure 6, where  $r_1=0.5$  inch,  $r_2=-0.838$  inch. For  $b=0.295$  inch, the surfaces are concentric. The refractive index  $N_D=1.523$ .

To vary the radius  $b$  of the zone from which light returns parallel to the incident beam, a special "button" whose length could be adjusted was made with oil of the same refractive index as the glass between the refracting and the reflecting surfaces. The intensity from this "button" was compared with the illumination at its surface by photographic photometry using the scheme shown in Figure 7, where  $S$  is a "Pointolite" lamp,  $L$  a condensing lens,  $F$  a filter giving approximately monochromatic light,  $F'$  nickel neutral filters to control

the exposure,  $A$  the source aperture,  $C$  a collimating lens of 72 inch focal length,  $R$  a parallel glass plate,  $O$  an objective of 18 inch focal length, and  $P$  the film. The camera was focused for infinity, converting the distribution of intensity with angle to a distribution of illumination across the focal plane. The "button"

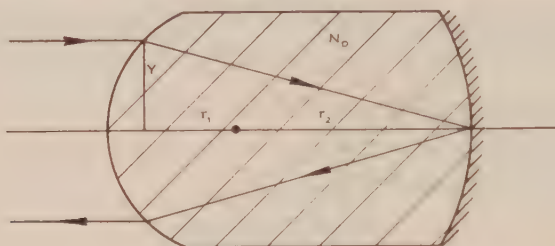


Fig. 6.—Optical system of reflector button.

was made about twice the size shown in Figure 6, and the results were corrected for the transmission of the oil and glass and the reflectivity of the silver, and then reduced to the scale of Figure 6. Comparison photographs of the illumination at the surface of the "button" were obtained by rotating the plate  $R$  through  $90^\circ$  and placing an aperture of known size in front of the camera objective  $O$ .

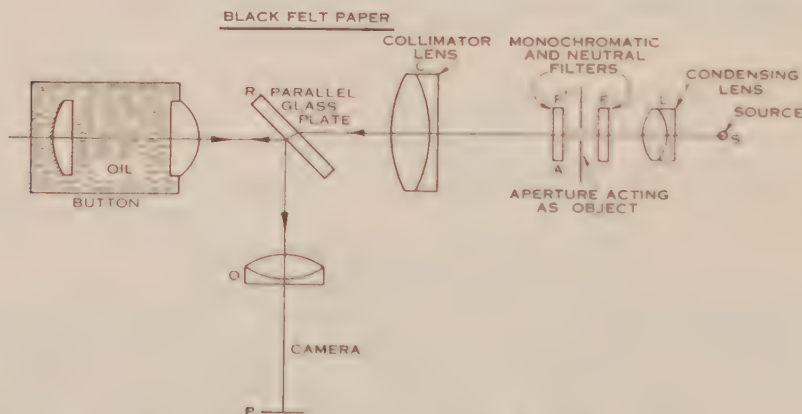


Fig. 7.—Experimental arrangement for intensity measurements.

In Plate 1 the distribution is shown for the "button" outside, at, and inside paraxial focus, while Plate 2 shows the zones on the "button" that send light back parallel to the axis, under the same conditions. For the "button" inside focus, further photographs in Plate 3 show the zones that send light along the optical axis, in the direction of the edge maximum intensity, and in two intermediate directions.

Ray traces were carried out for this optical system with various separations between the surfaces, and from these a table was made of  $a$  in terms of  $b$ , so that the algebraic approximation of the primary aberration treatment could be



applied. The intensity distribution was also calculated from the sets of numerical results from the ray traces using equation (22) and equations (13), (14), and (17, 18), omitting the cosines as the angles involved were small. In Figure 8

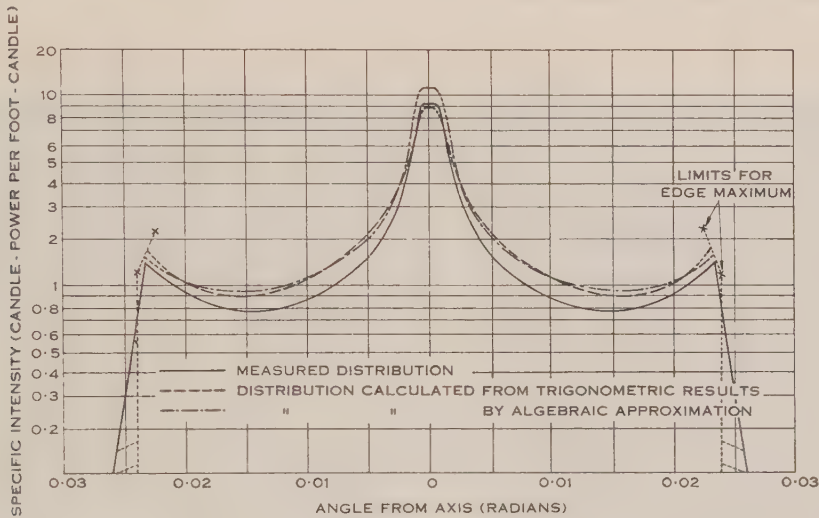


Fig. 8.—Light distribution from reflector button.

these two intensity distributions are compared with the measured results for one particular focus, while in Figure 9 the values of the central and the minimum specific intensity are shown for variation of  $a$ .

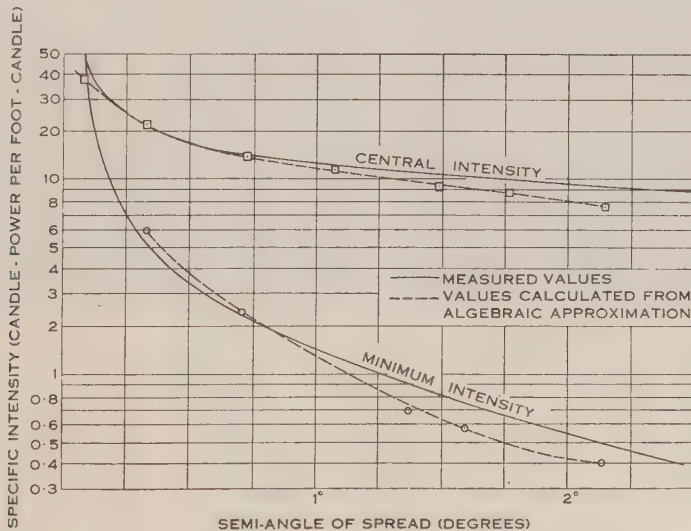


Fig. 9.—Reflector button—specific intensity *v.* angle.

The experimental results have random errors of about 5 per cent. and there may be systematic errors of about the same order. For the case shown in

Figure 8 and for other cases measured, the intensities calculated from the trigonometrical results agree very well at the centre of the distribution. The agreement is not so good at other angles where the areas that send light are not quite elliptical, as assumed.

The primary aberration treatment using the values of  $a$  and  $b$  found from ray traces does not give as good an approximation, but, as may be seen in Figure 9, it gives a fair estimate over a considerable range of focus.

## IX. DISCUSSION

The method of calculating light distributions, as outlined in this paper, is, within the limits of geometrical optics on which it is based, accurate for point sources, and gives an approximate solution for small finite sources. Referring again to the two classes into which the general problem can be divided, namely, the light distribution near or remote from a focus, in the former when diffraction is not negligible this method gives an approximate answer, while in the latter it provides a satisfactory solution for light projection systems of axial symmetry when the source is small.

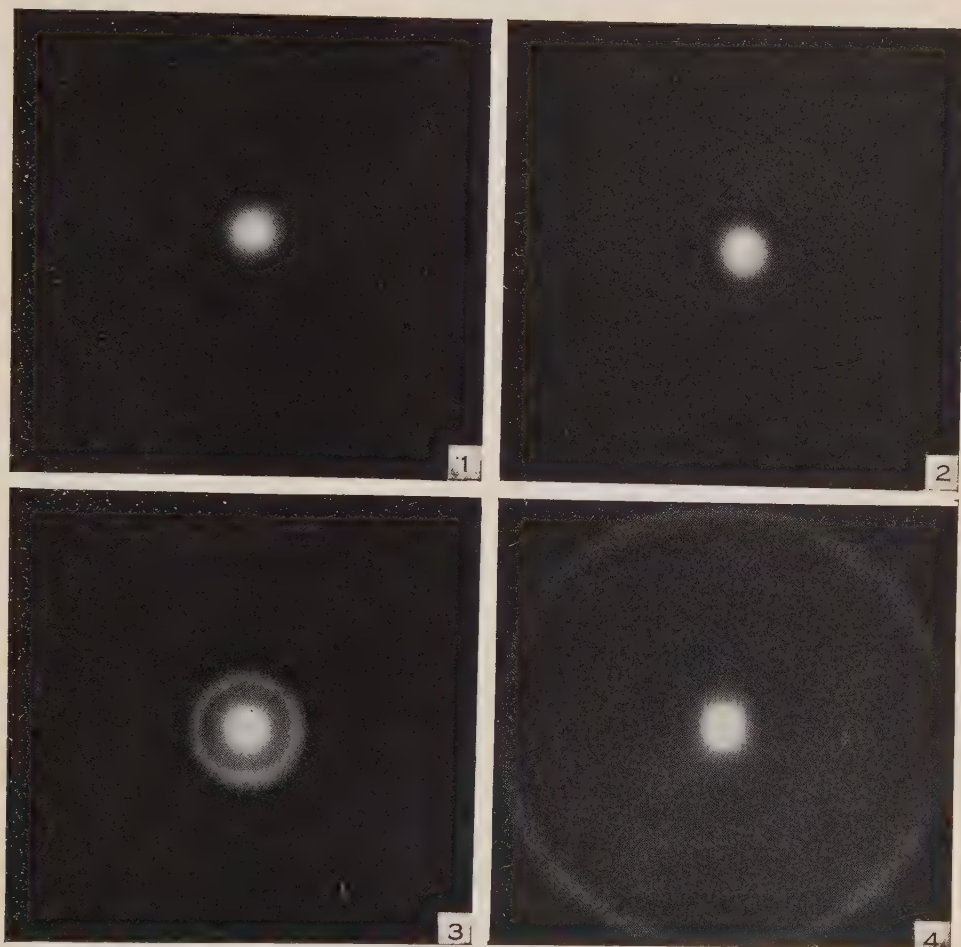
To apply the method to practical systems a set of rays from the centre of the object is traced through the system at various apertures and the required derivatives are obtained either by difference methods or by fitting to these results a polynomial for  $u'$  in terms of  $y$ . This latter method, and mechanical means of demonstrating the appearance of the distribution, have been given fully by van Heel(11). When new systems are being designed the simplified approximation, using powers up to  $y^3$  only, gives an algebraic result in which the effect of the parameters of the system can be seen directly.

## X. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Physics, C.S.I.R.O. The author wishes to thank Dr. R. G. Giovanelli, of that Division, for his help and criticism during the preparation of this paper, Miss J. Ward, who performed the calculations, and Miss C. Maguire, who assisted with the experimental work.

## XI. REFERENCES

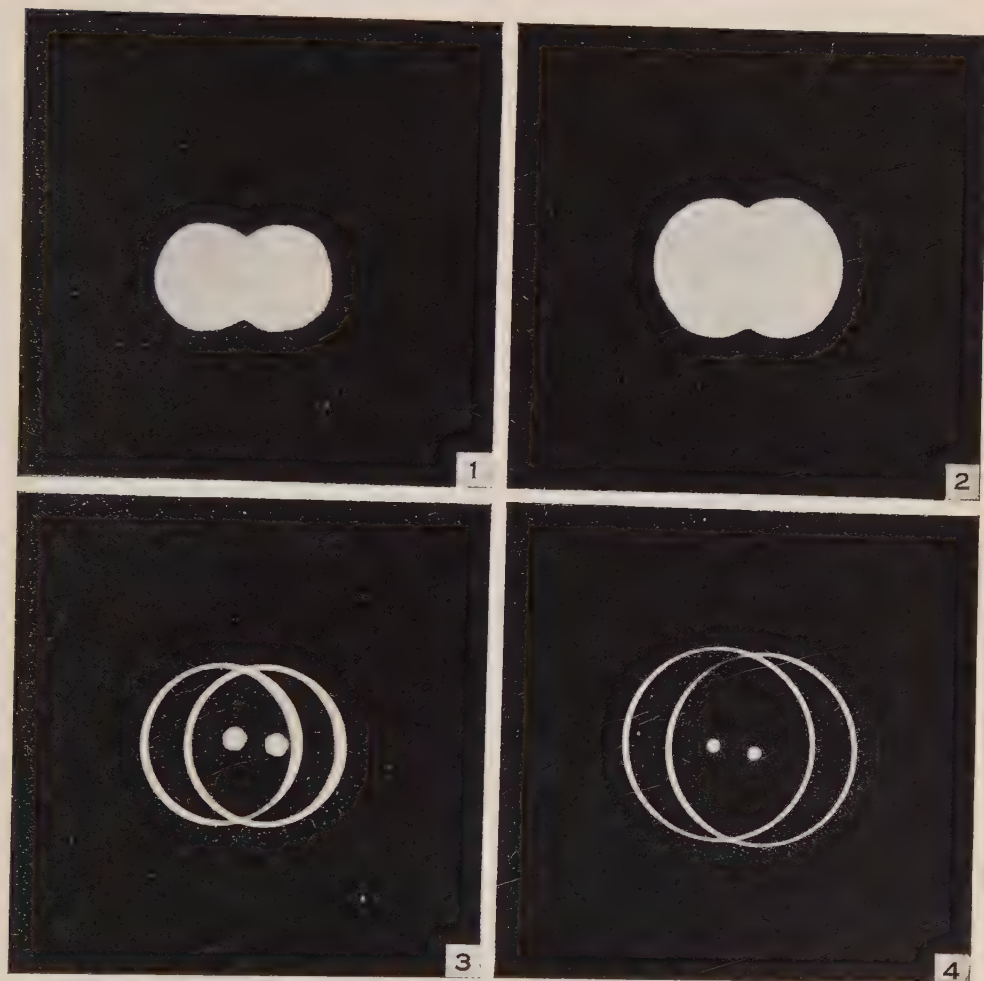
- (1) AIRY, G. B.—*Comb. Phil. Soc. Trans.* **5** : 238 (1834).
- (2) CONRADY, A. E.—*Mon. Not. R. Astr. Soc.* **79** : 575 (1919).
- (3) MARTIN, L. C.—*Trans. Opt. Soc.* **27** : 249 (1926).
- (4) MARÉCHAL, A.—*Rev. Opt. (Théor. Instrum.)* **26** : 257 (1947).
- (5) MARÉCHAL, A.—*Ibid.* **27** : 73 (1948).
- (6) HAWKINS, D., and LINFOOT, C. E.—*Mon. Not. R. Astr. Soc.* **105** : 334 (1945).
- (7) HERZBERGER, M.—*J. Opt. Soc. Amer.* **37** : 485 (1947).
- (8) BLONDEL, A., and LAVANCHY, C.—*Ann. Phys. Paris* **7** : 249 (1917).
- (9) BLONDEL, A., and LAVANCHY, C.—*Ibid.* **8** : 51 (1917).
- (10) DUNOYER, L.—*Rev. Opt. (Théor. Instrum.)* **27** : 399 (1948).
- (11) HEEL, A. C. S. VAN. "Achievements in Optics." Monograph on the Progress of Research in Holland during the War. p. 94. (Elsevier Publishing Co.: New York and Amsterdam, 1946.)



STEEL.—THE DISTRIBUTION OF LIGHT FROM OPTICAL SYSTEMS

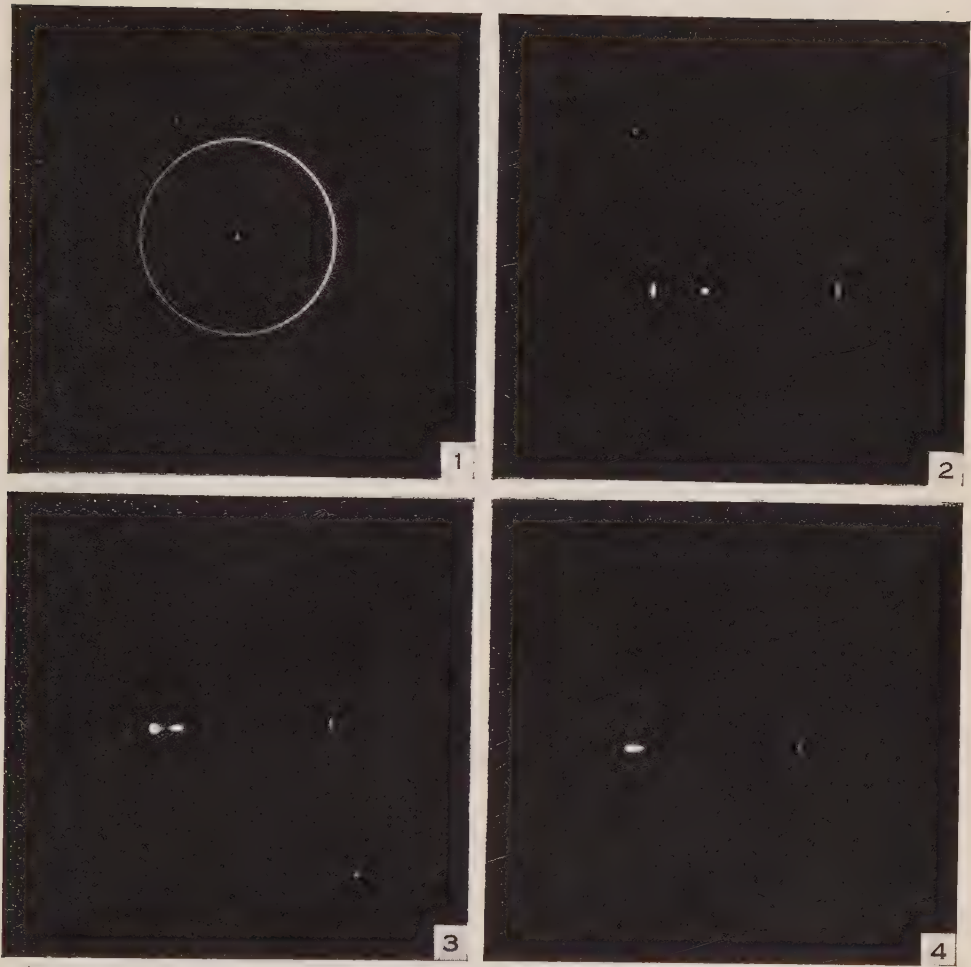






STEEL.—THE DISTRIBUTION OF LIGHT FROM OPTICAL SYSTEMS





STEEL.—THE DISTRIBUTION OF LIGHT FROM OPTICAL SYSTEMS





- (12) CONRADY, A. E.—“Applied Optics and Optical Design.” (Oxford Univ. Press : 1929.)
- (13) VAN LEAR, G. A.—*J. Opt. Soc. Amer.* **30** : 462 (1940).
- (14) SMITH, T.—“Dictionary of Applied Physics.” Vol. 4, p. 287. (Macmillan and Co. : London, 1923.)
- (15) SATCHE, P.—*Rev. Opt. (Théor. Instrum.)* **18** : 20 (1939).

## EXPLANATION OF PLATES 1-3

## PLATE 1

Figs. 1-4.—Light distribution from a reflector button outside, at, and inside paraxial focus.

## PLATE 2

Figs. 1-4.—Areas of the reflector button sending light parallel to the axis outside, at, and inside paraxial focus. Each image is doubled.

## PLATE 3

Figs. 1-4.—Areas of the button, focused to give the distribution of Plate 1, Figure 4, that send light along the axis and at various angles up to that of the edge maximum.

# THE UNIFORM TORSION OF AN INCOMPLETE TORE

By W. FREIBERGER\*

[Manuscript received December 8, 1948]

## Summary

A stress system is determined which gives as the resultant over any terminal cross-section a single force acting along the axis of the tore. This will yield a close approximation to the stresses in a helical spring of small pitch.

The method used is that of finding a stress function analogous to the ordinary torsion function subject to a certain boundary condition and the problem is solved in toroidal harmonics. The stresses and resultant force are evaluated in terms of the dimensions of the tore and compared with the results of known approximate theories.

## I. INTRODUCTION

When the terminal cross-sections of an incomplete tore are subjected to shearing actions with their resultant directed along the tore-axis, the system of stresses will approximate closely to that set up in a helical spring of small pitch under tension or compression. We present here an exact solution of this problem, in the sense that explicit expressions are obtained for the stresses at each point and for the resultant stress (i.e. the applied external force) in terms of very rapidly converging series.

The history of previous solutions to the problem was reviewed briefly by Southwell(1). Michell(2) obtained solutions for tores of approximately circular cross-section by basing his treatment on the displacement equilibrium equations and using a polynomial stress-function. Göhner(3) used successive approximations; his results were quoted by Timoshenko(4), together with a diagram exhibiting the distribution of shearing stress along the radius of a particular tore. Southwell(1) presented a formal solution for arbitrary cross-section with a view towards making the problem amenable to relaxation methods. All investigators have so far been working entirely in cylindrical polar coordinates.

The problem reduces to the determination of a single function which defines the stress-system. The partial differential equation satisfied by this stress-function is here solved in terms of toroidal harmonics because the use of a toroidal coordinate system enables us to satisfy the boundary conditions on the free surface of the tore. This leads to analytic expressions for the stress-system, which are evaluated for a special case; the stresses around the circumference and along the diameter of this special tore are shown graphically.

Great interest attaches to the manner in which the ratio of maximum stress to external force varies with the dimensions of the tore; this variation is plotted

\* Division of Aeronautics, C.S.I.R.

graphically in Figure 3, together with an approximate curve by Wahl(5) which only slightly overestimates the maximum stress. The solution is seen to converge to the well-known one for a circular cylinder when the thickness of the tore is indefinitely decreased.

## II. EQUATIONS OF EQUILIBRIUM. STRESS FUNCTION

We investigate the stress-system which is independent of angular distance along a generator of the tore (coordinate  $\theta$ ), the only non-vanishing stresses being the shearing actions  $\widehat{r\theta}$  and  $\widehat{\theta z}$  on meridian planes. The equations of equilibrium for this system reduce in cylindrical polar coordinates to the single equation (Southwell 1)

$$\frac{\partial}{\partial r}(r^2 \cdot \widehat{r\theta}) + \frac{\partial}{\partial z}(r^2 \cdot \widehat{\theta z}) = 0 \quad \dots\dots\dots (2.1)$$

and the Beltrami-Michell equations of compatibility, ensuring strains to be consistent with displacements, to

$$\left(\nabla^2 - \frac{4}{r^2}\right)\widehat{r\theta} = 0, \quad \left(\nabla^2 - \frac{1}{r^2}\right)\widehat{\theta z} = 0 \quad \dots\dots\dots (2.2)$$

Equation (2.1) shows that there exists a function  $\varphi$  such that

$$r^2 \cdot \widehat{r\theta} = A \frac{\partial \varphi}{\partial z}, \quad r^2 \cdot \widehat{\theta z} = -A \frac{\partial \varphi}{\partial r}, \quad \dots\dots\dots (2.3)$$

where  $A$  is a constant which affects the absolute values, but not the distribution of the stresses. Its value will later be related to that of the applied forces.

With (2.3) the equations (2.2) become

$$A \left(\nabla^2 - \frac{4}{r^2}\right) \frac{1}{r^2} \frac{\partial \varphi}{\partial z} = 0 \quad \dots\dots\dots (2.4)$$

and

$$A \left(\nabla^2 - \frac{1}{r^2}\right) \frac{1}{r^2} \frac{\partial \varphi}{\partial r} = 0 \quad \dots\dots\dots (2.5)$$

The operators  $\nabla^2$  and  $\partial/\partial z$  commute so that (2.4) immediately reduces to

$$A \frac{\partial}{\partial z} \left(\nabla^2 - \frac{4}{r^2}\right) \frac{\varphi}{r^2} = 0$$

or

$$A \left(\nabla^2 - \frac{4}{r^2}\right) \frac{\varphi}{r^2} = f_1(r), \quad \dots\dots\dots (2.6)$$

where  $f_1(r)$  is a function of  $r$  only.

The operators  $\nabla^2$  and  $\partial/\partial r$  do not commute, but after some reduction equation (2.5) becomes

$$A \frac{\partial}{\partial r} \left[ r^2 \left(\nabla^2 - \frac{4}{r^2}\right) \frac{\varphi}{r^2} \right] = 0$$

therefore

$$A \left(\nabla^2 - \frac{4}{r^2}\right) \frac{\varphi}{r^2} = \frac{f_2(z)}{r^2} \quad \dots\dots\dots (2.7)$$

Comparing (2.6) and (2.7) it follows that

$$r^2 f_1(r) = f_2(z) \dots\dots\dots (2.8)$$

The left-hand side of (2.8) is a function of  $r$  only, the right-hand side a function of  $z$  only, which is impossible unless each side equals a constant; this constant will be chosen as the  $A$  of (2.3).

Thus

$$f_1(r) = A/r^2, \quad f_2(z) = A \dots\dots\dots (2.9)$$

and each of (2.6) and (2.7) becomes

$$\left( \nabla^2 - \frac{4}{r^2} \right) \frac{\varphi}{r^2} = \frac{1}{r^2} \dots\dots\dots (2.10)$$

### III. TOROIDAL COORDINATES. BOUNDARY CONDITIONS

Toroidal coordinates  $(\sigma, \psi, \theta)$  will be introduced, related to orthogonal Cartesians  $(x, y, z)$ , and cylindrical polars  $(r, \theta, z)$  by the equations

$$\begin{aligned} x &= r \cos \theta, \quad y = r \sin \theta, \quad z + ir = ia \coth \frac{1}{2}(\sigma + i\psi) \\ r &= \frac{a \sinh \sigma}{\cosh \sigma - \cos \psi}, \quad z = \frac{a \sin \psi}{\cosh \sigma - \cos \psi} \dots\dots\dots (3.1) \end{aligned}$$

The surfaces  $\sigma = \text{constant}$  are a family of tores with the centre of their central lines at the origin 0, and their common axis along the  $z$ -direction.

The surfaces  $\psi = \text{constant}$  represent spherical bowls with their centres on the common axis of the tores  $\sigma = \text{constant}$ .

The toroidal coordinate system is produced by rotating the bipolar system of Figure 1 about the  $z$ -axis.

TABLE 1

DISTRIBUTION OF SHEAR STRESS ALONG A HORIZONTAL DIAMETER OF THE TORE  $R_0/\rho_0 = 4$ ; OF EQUATION (4.20) AND FIGURE 1

$\frac{R_0}{\rho_0}$	Shear Stress $\widehat{\psi\theta} = \lambda \frac{\mu c}{R_0}$	
	$\lambda$	
	$\psi = 0$	$\psi = 180^\circ$
4.0	0.1853	0.3500
4.1	0.1797	0.3417
4.5	0.1598	0.3131
5.0	0.1392	0.2846
6.0	0.1080	0.2432
8.0	0.06835	0.1936
10.0	0.04416	0.1649



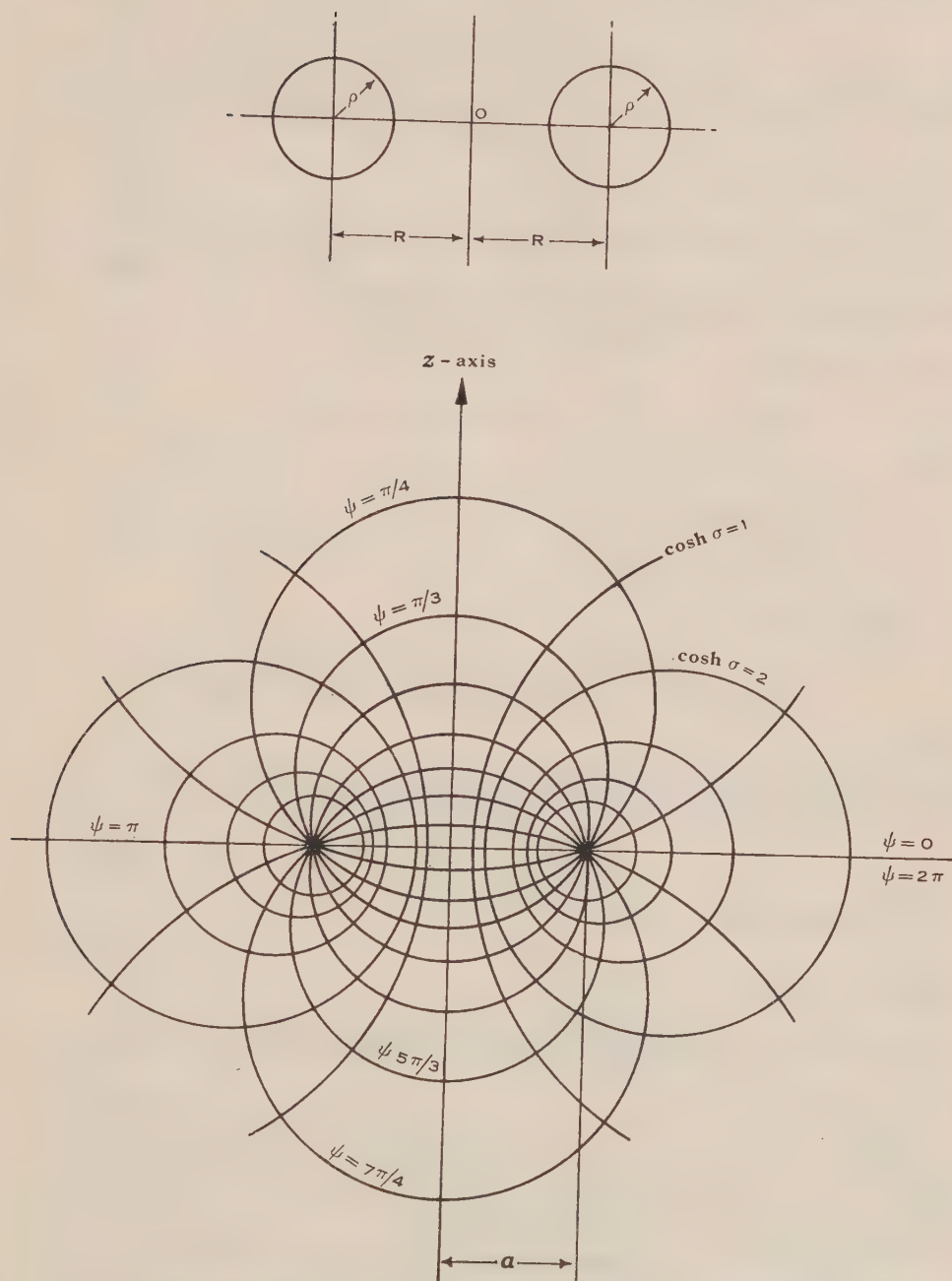


Fig. 1.— $R = a \frac{\cosh \sigma}{\sinh \sigma}$ ,  $\rho = \frac{a}{\sinh \sigma}$ .

By eliminating  $\psi$  from the equations (3.1) we obtain

$$(r-a \coth \sigma)^2 + z^2 = a^2 \operatorname{cosech}^2 \sigma \dots\dots\dots (3.2)$$

which represents the equation of a circle of radius  $\rho = a \operatorname{cosech} \sigma$ , centre at  $z=0$ ,  $r=a \coth \sigma$  ( $=R$ , say). The radius of the circle of centroids, which may be called the central line of the tore  $\sigma=\text{constant}$ , is therefore

$$R = a \coth \sigma \dots\dots\dots (3.3)$$

and the radius of a circular cross-section

$$\rho = a \operatorname{cosech} \sigma; \dots\dots\dots (3.4)$$

this gives for the quantity characterizing the dimensions of the tore

$$R/\rho = \cosh \sigma \dots\dots\dots (3.5)$$

The line element in orthogonal curvilinear coordinates  $\alpha_i$

$$ds^2 = h_1^2 d\alpha_1^2 + h_2^2 d\alpha_2^2 + h_3^2 d\alpha_3^2 \dots\dots\dots (3.6)$$

becomes in toroidals

$$ds^2 = \frac{a^2}{(\cosh \sigma - \cos \psi)^2} (d\sigma^2 + d\psi^2 + \sinh^2 \sigma d\theta^2) \dots\dots\dots (3.7)$$

so that

$$\left. \begin{aligned} h_1 = h_2 &= \frac{a}{\cosh \sigma - \cos \psi} = h, \text{ say,} \\ h_3 &= \frac{a \sinh \sigma}{\cosh \sigma - \cos \psi}. \end{aligned} \right\} \dots\dots\dots (3.8)$$

The Laplacian operator

$$\nabla^2 = \frac{1}{h_1 h_2 h_3} \left[ \frac{\partial}{\partial \alpha_1} \left( \frac{h_2 h_3}{h_1} \frac{\partial}{\partial \alpha_1} \right) + \frac{\partial}{\partial \alpha_2} \left( \frac{h_3 h_1}{h_2} \frac{\partial}{\partial \alpha_2} \right) + \frac{\partial}{\partial \alpha_3} \left( \frac{h_1 h_2}{h_3} \frac{\partial}{\partial \alpha_3} \right) \right]$$

transforms into

$$\nabla^2 = \frac{(\cosh \sigma - \cos \psi)^3}{a^3 \sinh \sigma} \left[ \frac{\partial}{\partial \sigma} \left( \frac{a \sinh \sigma}{\cosh \sigma - \cos \psi} \frac{\partial}{\partial \sigma} \right) + \frac{\partial}{\partial \psi} \left( \frac{a \sinh \sigma}{\cosh \sigma - \cos \psi} \frac{\partial}{\partial \psi} \right) + \frac{a}{(\cosh \sigma - \cos \psi) \sinh \sigma} \frac{\partial^2}{\partial \theta^2} \right] \dots\dots (3.9)$$

so that equation (2.10) with

$$V = \varphi/r^2 \dots\dots\dots (3.10)$$

say, and  $V$  independent of  $\theta$  becomes

$$(\cosh \sigma - \cos \psi) \sinh \sigma \left[ \frac{\partial}{\partial \sigma} \left( \frac{\sinh \sigma}{\cosh \sigma - \cos \psi} \frac{\partial V}{\partial \sigma} \right) + \frac{\partial}{\partial \psi} \left( \frac{\sinh \sigma}{\cosh \sigma - \cos \psi} \frac{\partial V}{\partial \psi} \right) \right] - 4V = 1 \dots\dots\dots (3.11)$$

Before proceeding to solve this equation, we have to express the boundary conditions in toroidal coordinates; to this end we transform the components of stress from the polar coordinates into this system and seek the relation which will then connect them with the stress function.

It is obvious from symmetry that all stress components apart from  $\sigma_\theta$  and  $\psi_\theta$  vanish identically. From the formulae for the transformation of stresses under rotation of the coordinate system

$$\left. \begin{aligned} \widehat{\sigma_\theta} &= \widehat{\theta z} \sin \gamma + r \widehat{\theta} \cos \gamma \\ \widehat{\psi_\theta} &= -\widehat{\theta z} \cos \gamma + r \widehat{\theta} \sin \gamma \end{aligned} \right\} \dots\dots\dots (3.12)$$

An element of length in the direction of the  $r$ -coordinates is given by  $h_r dr = dr$ , since  $h_r = 1$  follows from

$$\begin{aligned} ds^2 &= dr^2 + r^2 d\theta^2 + dz^2 \\ &= h_r dr^2 + h_\theta d\theta^2 + h_z dz^2, \end{aligned}$$

and an element in the  $\sigma$ -direction by

$$h_\sigma d\sigma = h_1 d\sigma = h d\sigma \quad (\text{cf. (3.6), (3.7)})$$

so that, from Figure 2,

$$\cos \gamma = \frac{h_\sigma d\sigma}{h_r dr} = h \frac{\partial \sigma}{\partial r} \dots \dots \dots (3.13)$$

It follows similarly that

$$\cos \left( \frac{\pi}{2} - \gamma \right) = \sin \gamma = \frac{h_\sigma d\sigma}{h_z dz} = h \frac{\partial \sigma}{\partial z} \dots \dots \dots (3.14)$$



Fig. 2

By considering the angle  $\gamma$  as extending on the one hand positively from the  $\sigma$ - to the  $r$ -direction, on the other from the  $r$ - to the  $\sigma$ -direction, and equating the cosines of the two cases, it is clear that

$$h \frac{\partial \sigma}{\partial r} = \frac{1}{h} \frac{\partial r}{\partial \sigma} \dots \dots \dots (3.15)$$

and similarly

$$h \frac{\partial \sigma}{\partial z} = \frac{1}{h} \frac{\partial z}{\partial \sigma} \dots \dots \dots (3.16)$$

The Cauchy-Riemann differential equations for the complex function

$$z + ir = ia \coth \frac{1}{2}(\sigma + i\psi)$$

are

$$\frac{\partial z}{\partial \sigma} = \frac{\partial r}{\partial \psi} \dots \dots \dots (3.17)$$

and

$$\frac{\partial z}{\partial \psi} = -\frac{\partial r}{\partial \sigma} \dots \dots \dots (3.18)$$

We can now express  $\widehat{\sigma\theta}$ ,  $\widehat{\psi\theta}$  in terms of  $\varphi$ . Substituting (3.13) and (3.14) in (3.12) we get

$$\widehat{\sigma\theta} = \widehat{\theta z} \cdot h \frac{\partial \sigma}{\partial z} + r \widehat{\theta} \cdot h \frac{\partial \sigma}{\partial r}$$

$$= \widehat{\theta} z \cdot \frac{1}{h} \frac{\partial r}{\partial \psi} - r \widehat{\theta} \cdot \frac{1}{h} \frac{\partial z}{\partial \psi} \quad \text{from (3.15) to (3.18)}$$

$$= -\frac{A}{r^2 h} \left( \frac{\partial \varphi}{\partial r} \frac{\partial r}{\partial \psi} + \frac{\partial \varphi}{\partial z} \frac{\partial z}{\partial \psi} \right) \quad \text{from (2.3)}$$

so that

$$\sigma \widehat{\theta} = -\frac{A}{r^2 h} \frac{\partial \varphi}{\partial \psi} \dots \dots \dots (3.19)$$

Similarly

$$\widehat{\psi} \theta = -\widehat{\theta} z \cdot h \frac{\partial \sigma}{\partial r} + r \widehat{\theta} \cdot h \frac{\partial \sigma}{\partial z}$$

$$= -\widehat{\theta} z \cdot \frac{1}{h} \frac{\partial r}{\partial \sigma} + r \widehat{\theta} \cdot \frac{1}{h} \frac{\partial z}{\partial \sigma}$$

$$= \frac{A}{r^2 h} \left( \frac{\partial \varphi}{\partial r} \frac{\partial r}{\partial \sigma} + \frac{\partial \varphi}{\partial z} \frac{\partial z}{\partial \sigma} \right)$$

so that

$$\widehat{\psi} \theta = \frac{A}{r^2 h} \frac{\partial \varphi}{\partial \sigma} \dots \dots \dots (3.20)$$

The boundary condition on the free surface of the tore is that the normal stress  $\widehat{\sigma} \theta$  be made to vanish there. In terms of the stress function this means, from (3.19), that

$$\frac{\partial \varphi}{\partial \psi} = 0 \quad \text{on the boundary} \dots \dots \dots (3.21)$$

that is

$$\varphi = \text{constant} \dots \dots \dots (3.22)$$

Since the addition of a constant value to  $\varphi$  does not affect the stress system, (3.22) can be replaced by

$$\varphi = 0 \quad \text{on the boundary } \sigma = \sigma_0, \text{ say} \dots \dots (3.23)$$

#### IV. SOLUTION OF THE DIFFERENTIAL EQUATION

The general solution of equation (3.11) is given by the sum of the general solution of the homogeneous equation

$$(\cosh \sigma - \cos \psi) \sinh \sigma \left[ \frac{\partial}{\partial \sigma} \left( \frac{\sinh \sigma}{\cosh \sigma - \cos \psi} \frac{\partial V}{\partial \sigma} \right) + \frac{\partial}{\partial \psi} \left( \frac{\sinh \sigma}{\cosh \sigma - \cos \psi} \frac{\partial V}{\partial \psi} \right) \right] - 4V = 0 \dots \dots \dots (4.1)$$

and a particular solution of (3.11).

An obvious particular solution of (3.11) is

$$V = -\frac{1}{4} \dots \dots \dots (4.2)$$

so that we concentrate on the general solution of (4.1). In order to bring it into an integrable form we put

$$V = U (\cosh \sigma - \cos \psi)^{\frac{1}{2}} \dots \dots \dots (4.3)$$

After some reduction the equation becomes

$$\frac{\partial^2 U}{\partial \sigma^2} + \frac{\partial^2 U}{\partial \psi^2} + \coth \sigma \frac{\partial U}{\partial \sigma} + \left( \frac{1}{4} - \frac{4}{\sinh^2 \sigma} \right) U = 0 \dots \dots \dots (4.4)$$



Writing

$$U = H(\sigma) \cdot \psi(\psi) \dots \dots \dots (4.5)$$

to separate the variables and dividing (4.4) by  $U$  gives

$$\frac{1}{H} \frac{d^2 H}{d\sigma^2} + \frac{1}{\psi} \frac{d^2 \psi}{d\psi^2} + \frac{\coth \sigma}{H} \frac{dH}{d\sigma} + \frac{1}{4} - \frac{4}{\sinh^2 \sigma} = 0$$

that is

$$\frac{1}{H} \frac{d^2 H}{d\sigma^2} + \frac{\coth \sigma}{H} \frac{dH}{d\sigma} + \frac{1}{4} - \frac{4}{\sinh^2 \sigma} = -\frac{1}{\psi} \frac{d^2 \psi}{d\psi^2} \dots \dots (4.6)$$

Each side of (4.6) is dependent on a different variable and so must have a common constant value  $n^2$ , say; hence

$$\frac{d^2 \psi}{d\psi^2} + n^2 \psi = 0$$

therefore

$$\psi = \alpha \cos n\psi + \beta \sin n\psi, \dots \dots \dots (4.7)$$

where  $\alpha, \beta$  are arbitrary constants.

The left-hand side of (4.6) becomes

$$\frac{d^2 H}{d\sigma^2} + \coth \sigma \frac{dH}{d\sigma} - \left( n^2 - \frac{1}{4} + \frac{4}{\sinh^2 \sigma} \right) H = 0 \dots \dots \dots (4.8)$$

This equation is a special case of the differential equation (Whittaker and Watson 6, p. 324)

$$(1 - \xi^2) \frac{d^2 w}{d\xi^2} - 2\xi \frac{dw}{d\xi} + \left[ n'(n'+1) - \frac{m^2}{1 - \xi^2} \right] w = 0 \dots \dots (4.9)$$

which is satisfied by Legendre's associated functions of the first and second kind

$$P_{n'}^m(\xi) \text{ and } Q_{n'}^m(\xi).$$

By substituting in (4.9)  $\xi = \cosh \sigma$ ,

$$\frac{d^2 w}{d\sigma^2} + \coth \sigma \frac{dw}{d\sigma} - \left[ n'(n'+1) + \frac{m^2}{\sinh^2 \sigma} \right] w = 0 \dots \dots (4.10)$$

Comparing (4.8) with (4.10) shows that  $H(\sigma)$  satisfies equation (4.10) provided

$$n' = n - \frac{1}{2}, \quad m = \pm 2; \dots \dots \dots (4.11)$$

so, since  $m = +2$  and  $m = -2$  give linearly dependent solutions,  $H(\sigma)$  has the form of linear combinations of

$$P_{n-\frac{1}{2}}^2(\cosh \sigma) \text{ and } Q_{n-\frac{1}{2}}^2(\cosh \sigma).$$

The stresses, and so the stress function, must be single-valued: it follows from (4.7), therefore, that  $\psi$  has to be periodic of period  $2\pi$  and consequently  $n$  integral.

In problems involving solid tori the associated Legendre functions of the first kind are inapplicable since they have the asymptotic value

$$P_{n-\frac{1}{2}}^m(\cosh \sigma) \sim e^{(n+\frac{1}{2})\sigma} \times \text{constant}$$

for large  $\sigma$  (Hobson 7, p. 436) and this increases indefinitely as  $\sigma \rightarrow \infty$ .

The equation (4.4) therefore has the solution

$$U = \sum_{n=0}^{\infty} Q_{n-\frac{1}{2}}^2 (\cosh \sigma) (A_n \cos n\psi + B_n \sin n\psi)$$

and so (4.1):

$$V = (\cosh \sigma - \cos \psi)^{\frac{1}{2}} \sum_{n=0}^{\infty} Q_{n-\frac{1}{2}}^2 (\cosh \sigma) (A_n \cos n\psi + B_n \sin n\psi) \dots (4.12)$$

Adding to this the special solution (4.2) and remembering the definition (3.10), the expression for the stress function becomes

$$\varphi(\sigma, \psi) = \frac{a^2 \sinh^2 \sigma}{(\cosh \sigma - \cos \psi)^2} \left[ (\cosh \sigma - \cos \psi)^{\frac{1}{2}} \sum_{n=0}^{\infty} Q_{n-\frac{1}{2}}^2 (\cosh \sigma) (A_n \cos n\psi + B_n \sin n\psi) - \frac{1}{4} \right] \dots (4.13)$$

The coefficients  $A_n, B_n$  must now be chosen in such a manner that the boundary condition (3.23) on the free surface of the tore be satisfied. Substituting (4.13) in (3.23) requires

$$(\cosh \sigma_0 - \cos \psi)^{\frac{1}{2}} \sum_{n=0}^{\infty} Q_{n-\frac{1}{2}}^2 (\cosh \sigma_0) (A_n \cos n\psi + B_n \sin n\psi) = \frac{1}{4}$$

that is

$$\sum_{n=0}^{\infty} Q_{n-\frac{1}{2}}^2 (\cosh \sigma_0) (A_n \cos n\psi + B_n \sin n\psi) = \frac{1}{4} (\cosh \sigma_0 - \cos \psi)^{-\frac{1}{2}} \dots (4.14)$$

The right-hand side is an even function of  $\psi$ , so that on the left-hand side  $B_n$  must be identically zero for all  $n$ .

Expanding the function

$$f(\psi) = (\cosh \sigma_0 - \cos \psi)^{-\frac{1}{2}}$$

in a Fourier cosine series

$$\frac{1}{2}a_0 + a_1 \cos \psi + a_2 \cos 2\psi + \dots + a_n \cos n\psi + \dots (4.15)$$

we have

$$a_n = \frac{2}{\pi} \int_0^\pi f(\psi) \cos n\psi d\psi = \frac{2}{\pi} \int_0^\pi \frac{\cos n\psi d\psi}{(\cosh \sigma_0 - \cos \psi)^{\frac{1}{2}}} \dots (4.16)$$

$$(n=0, 1, 2, \dots)$$

Substituting this series in the right-hand side of (4.14) and comparing coefficients of  $\cos n\psi$  for  $n=0, 1, 2, \dots$  gives

$$\left. \begin{aligned} A_0 &= \frac{1}{4\pi Q_{-\frac{1}{2}}^2 (\cosh \sigma_0)} \int_0^\pi \frac{d\psi}{(\cosh \sigma_0 - \cos \psi)^{\frac{1}{2}}} \\ A_k &= \frac{1}{2\pi Q_{k-\frac{1}{2}}^2 (\cosh \sigma_0)} \int_0^\pi \frac{\cos k\psi d\psi}{(\cosh \sigma_0 - \cos \psi)^{\frac{1}{2}}} \end{aligned} \right\} \dots (4.17)$$

( $k=1, 2, 3, \dots$ )

By an integral formula for Legendre functions of the second kind (Bateman 8, p. 464)

$$\int_0^\pi \frac{\cos n\psi d\psi}{(\cosh \sigma_0 - \cos \psi)^{\frac{1}{2}}} = \sqrt{2} Q_{n-\frac{1}{2}} (\cosh \sigma_0) \quad (n=0, 1, 2, \dots)$$

so that in equation (4.13)

$$\left. \begin{aligned} A_0 &= \frac{1}{2\pi\sqrt{2}} \frac{Q_{-\frac{1}{2}}(\cosh \sigma_0)}{Q_{-\frac{1}{2}}^2(\cosh \sigma_0)} \\ A_k &= \frac{1}{\pi\sqrt{2}} \frac{Q_{k-\frac{1}{2}}(\cosh \sigma_0)}{Q_{k-\frac{1}{2}}^2(\cosh \sigma_0)} \quad (k=1, 2, 3, \dots) \end{aligned} \right\} \dots (4.18)$$

and  $B_n=0$  ( $n=1, 2, \dots$ )

The stresses  $\widehat{\sigma\theta}$  and  $\widehat{\psi\theta}$  at any point  $(\sigma, \psi)$  of a tore with bounding surface  $\sigma=\sigma_0$  are therefore obtained by substituting in (3.19) and (3.20) for  $\varphi(\sigma, \psi)$  from (4.13) with the coefficients  $A_n, B_n$  as given by (4.18). The resulting expressions are

$$\frac{a}{A} \widehat{\sigma\theta} = (\cosh \sigma - \cos \psi)^{3/2} \sum_{n=0}^{\infty} \left\{ \left( \frac{3 \sin \psi \cos n\psi}{2 (\cosh \sigma - \cos \psi)} + n \sin n\psi \right) A_n Q_{n-\frac{1}{2}}^2(\cosh \sigma) \right\} \\ - \frac{1}{2} \sin \psi \dots \dots \dots (4.19)$$

$$\frac{a}{A} \widehat{\psi\theta} = \frac{(\cosh \sigma - \cos \psi)^{3/2}}{\sinh \sigma} \sum_{n=0}^{\infty} \left\{ \left( \frac{\cosh \sigma (\cosh \sigma - 4 \cos \psi) + 3}{2 (\cosh \sigma - \cos \psi)} Q_{n-\frac{1}{2}}^2(\cosh \sigma) \right. \right. \\ \left. \left. + \sinh^2 \sigma \frac{dQ_{n-\frac{1}{2}}^2(\cosh \sigma)}{d(\cosh \sigma)} \right) A_n \cos n\psi + \frac{\cosh \sigma \cos \psi - 1}{2 \sinh \sigma} \right\} \dots \dots (4.20)$$

An independent verification of the stress system (4.19)-(4.20) is provided by demonstrating that it satisfies the stress equations of equilibrium in toroidal coordinates; the only non-vanishing equation is (Sokolnikoff 9, p. 201)

$$\frac{\partial}{\partial \sigma} (h h_3 \widehat{\sigma\theta}) + \frac{\partial}{\partial \psi} (h h_3 \widehat{\psi\theta}) + \widehat{\sigma\theta} h \frac{\partial h_3}{\partial \sigma} + \widehat{\psi\theta} h \frac{\partial h_3}{\partial \psi} = 0$$

and this is easily shown to be true for  $h, h_3$  from (3.8),  $\widehat{\sigma\theta}, \widehat{\psi\theta}$  from (4.19) and (4.20).

TABLE 2

DISTRIBUTION OF SHEAR STRESS AROUND THE CIRCUMFERENCE OF THE TORE  $R_0/\rho_0=4$ ; CF. EQUATION (4.20) AND FIGURE 2

Shear Stress  $\widehat{\psi\theta} = \lambda \frac{\mu c}{R_0}$

$\psi$	$-\lambda$
$0^\circ$	0.1853
$30^\circ$	0.1959
$60^\circ$	0.2249
$90^\circ$	0.2657
$120^\circ$	0.3074
$150^\circ$	0.3385
$180^\circ$	0.3500

## V. RESULTANT STRESSES

It remains now to relate the undetermined constant  $A$  in (4.19)-(4.20) to the applied external force, and to show that our stress system has a resultant which is directed along the axis of the toro (the  $z$ -direction).

The resultant in the  $r$ -direction of the stresses acting on a cross-section is given by

$$\int \int \widehat{r\theta} \, dr dz = A \int \int \frac{\partial}{\partial z} \left( \frac{\varphi}{r^2} \right) dr dz = A \int \frac{dr}{ds} \frac{\varphi}{r^2} ds = 0$$

$c$

since  $\varphi=0$  on the boundary.

The moment of the stresses about the origin is

$$\begin{aligned} \int \int (z \widehat{r\theta} - r \widehat{\theta z}) dr dz &= A \int \int \left\{ z \frac{\partial}{\partial z} \left( \frac{\varphi}{r^2} \right) - \frac{1}{r} \frac{\partial \varphi}{\partial r} \right\} dr dz \\ &= A \int \int \left\{ \frac{\partial}{\partial r} \left( \frac{\varphi}{r} \right) + \frac{\partial}{\partial z} \left( \frac{z\varphi}{r^2} \right) \right\} dr dz \\ &= A \int \left( r \frac{dz}{ds} - z \frac{dr}{ds} \right) \frac{\varphi}{r^2} ds \quad (\text{by Stokes's theorem}) \\ &= 0 \quad \text{for the same reason.} \end{aligned}$$

The only non-vanishing resultant stress across any meridian section is therefore a single force  $Z$  in the axis of the toro; this is given by

$$Z = \int \int \widehat{\theta z} \, dr dz = - \int \int \frac{A}{r^2} \frac{\partial \varphi}{\partial r} dr dz \quad \dots (5.1)$$

From (2.10)

$$\left( \frac{\partial}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{\partial^2}{\partial z^2} - \frac{4}{r^2} \right) \frac{\varphi}{r^2} = \frac{1}{r^2},$$

and it follows after some manipulation that

$$\frac{\partial^2 \varphi}{\partial r^2} - \frac{3}{r} \frac{\partial \varphi}{\partial r} + \frac{\partial^2 \varphi}{\partial z^2} = 1;$$

hence

$$\frac{\partial}{\partial r} \left( \frac{1}{r} \frac{\partial \varphi}{\partial r} \right) - \frac{2}{r^2} \frac{\partial \varphi}{\partial r} + \frac{\partial}{\partial z} \left( \frac{1}{r} \frac{\partial \varphi}{\partial z} \right) = \frac{1}{r}$$

therefore

$$-\frac{1}{r^2} \frac{\partial \varphi}{\partial r} = \frac{1}{2} \left[ \frac{1}{r} - \frac{\partial}{\partial r} \left( \frac{1}{r} \frac{\partial \varphi}{\partial r} \right) - \frac{\partial}{\partial z} \left( \frac{1}{r} \frac{\partial \varphi}{\partial z} \right) \right].$$

Substituting this expression in (5.1) we obtain

$$\begin{aligned} \frac{Z}{A} &= \frac{1}{2} \int \int \left\{ \frac{1}{r} - \frac{\partial}{\partial r} \left( \frac{1}{r} \frac{\partial \varphi}{\partial r} \right) - \frac{\partial}{\partial z} \left( \frac{1}{r} \frac{\partial \varphi}{\partial z} \right) \right\} dr dz \\ &= \frac{1}{2} \int \int \left\{ \frac{\partial}{\partial r} \left( \log r - \frac{1}{r} \frac{\partial \varphi}{\partial r} \right) - \frac{\partial}{\partial z} \left( \frac{1}{r} \frac{\partial \varphi}{\partial z} \right) \right\} dr dz \\ &= \frac{1}{2} \int_0^{2\pi} \left\{ \frac{1}{r} \frac{\partial \varphi}{\partial z} \frac{\partial r}{\partial \psi} + \left( \log r - \frac{1}{r} \frac{\partial \varphi}{\partial r} \right) \frac{\partial z}{\partial \psi} \right\} d\psi \quad (\text{by Stokes's theorem}) \end{aligned}$$



$$\begin{aligned}
 &= \frac{1}{2} \int_0^{2\pi} \left\{ -\log r \frac{\partial r}{\partial \sigma} + \frac{1}{r} \frac{\partial \varphi}{\partial z} \frac{\partial z}{\partial \sigma} + \frac{\partial \varphi}{\partial r} \frac{\partial r}{\partial \sigma} \right\} d\psi \\
 &= \frac{1}{2} \int_0^{2\pi} \left( -\frac{\partial r}{\partial \sigma} \log r + \frac{1}{r} \frac{\partial \varphi}{\partial \sigma} \right) d\psi \dots\dots\dots (5.2)
 \end{aligned}$$

where the integration is along the boundary of a cross-section  $\theta = \text{constant}$ ,  $\psi$  changing from 0 to  $2\pi$ ; along this boundary we also have  $\sigma = \text{constant} = \sigma_0$ , so that the integrand of (5.2) must be evaluated for this value of  $\sigma$ . It is useful to remember that  $\varphi(\sigma, \psi) = 0$  for  $\sigma = \sigma_0$ .

Substituting for  $r, \frac{\partial r}{\partial \sigma}$  from Section III in (5.2) we obtain

$$\begin{aligned}
 \frac{Z}{A} &= \frac{a}{2} \int_0^{2\pi} \frac{1 - \cosh \sigma_0 \cos \psi}{(\cosh \sigma_0 - \cos \psi)^2} \log \frac{a \sinh \sigma_0}{\cosh \sigma_0 - \cos \psi} d\psi \\
 &\quad - \frac{1}{2a} \int_0^{2\pi} \frac{\cosh \sigma_0 - \cos \psi}{\sinh \sigma_0} \frac{\partial \varphi}{\partial \sigma} \bigg|_{\sigma=\sigma_0} d\psi \\
 &= I_a + I_b, \text{ say } \dots\dots\dots (5.3)
 \end{aligned}$$

Now

$$\begin{aligned}
 I_a &= a \int_0^\pi \frac{1 - \cosh \sigma_0 \cos \psi}{(\cosh \sigma_0 - \cos \psi)^2} \log (a \sinh \sigma_0) d\psi \\
 &\quad - a \int_0^\pi \frac{1 - \cosh \sigma_0 \cos \psi}{(\cosh \sigma_0 - \cos \psi)} \log (\cosh \sigma_0 - \cos \psi) d\psi \\
 &= -a \log (a \sinh \sigma_0) \left[ \frac{\sin \psi}{\cosh \sigma_0 - \cos \psi} \right]_0^\pi \\
 &\quad - a \left\{ \int_0^\pi \frac{\sin^2 \psi}{(\cosh \sigma_0 - \cos \psi)^2} - \left[ \frac{\sin \psi}{\cosh \sigma_0 - \cos \psi} \log (\cosh \sigma_0 - \cos \psi) \right]_0^\pi \right\}
 \end{aligned}$$

since

$$\begin{aligned}
 \frac{d}{d\psi} \left\{ \frac{\sin \psi}{\cosh \sigma - \cos \psi} \log (\cosh \sigma - \cos \psi) \right\} &= \frac{\cosh \sigma \cos \psi - 1}{(\cosh \sigma - \cos \psi)^2} \log (\cosh \sigma - \cos \psi) \\
 &\quad + \frac{\sin^2 \psi}{(\cosh \sigma - \cos \psi)^2}.
 \end{aligned}$$

Hence

$$\begin{aligned}
 I_a &= -a \int_0^\pi \frac{\sin^2 \psi d\psi}{(\cosh \sigma_0 - \cos \psi)^2} = -a \int_0^\pi \frac{1 - \cos^2 \psi}{(\cosh \sigma_0 - \cos \psi)^2} d\psi \\
 &= -a \int_0^\pi \frac{-1 + 2 + \cosh^2 \sigma - 2 \cosh \sigma \cos \psi - (\cosh \sigma - \cos \psi)^2}{(\cosh \sigma_0 - \cos \psi)^2} d\psi \\
 &= a \int_0^\pi d\psi - a \int_0^\pi \frac{\sinh^2 \sigma_0}{(\cosh \sigma_0 - \cos \psi)^2} d\psi + 2a \int_0^\pi \frac{\cosh \sigma_0 - \cos \psi - 1}{(\cosh \sigma_0 - \cos \psi)^2} d\psi
 \end{aligned}$$

$$\begin{aligned}
 &= a\pi - a \sinh^2 \sigma_0 \int_0^\pi \frac{d\psi}{(\cosh \sigma_0 - \cos \psi)^2} \\
 &= a\pi - a \sinh^2 \sigma_0 \cdot \frac{\pi \cosh \sigma_0}{\sinh^3 \sigma_0}
 \end{aligned}$$

therefore

$$I_a = a\pi (1 - \coth \sigma_0) \dots \dots \dots (5.4)$$

To evaluate  $I_b$  we require  $\partial\varphi/\partial\sigma$  at  $\sigma=\sigma_0$ ; from (4.13)

$$\varphi(\sigma, \psi) = \frac{a^2 \sinh^2 \sigma}{(\cosh \sigma - \cos \psi)^2} \left[ (\cosh \sigma - \cos \psi)^{\frac{1}{2}} \sum_{n=0}^{\infty} Q_{n-\frac{1}{2}}^2 (\cosh \sigma) A_n \cos n\psi - \frac{1}{4} \right],$$

where  $A_n$  is given by (4.18). Hence, at  $\sigma=\sigma_0$

$$\begin{aligned}
 \left. \frac{\partial\varphi}{\partial\sigma} \right|_{\sigma=\sigma_0} &= a^2 \sinh^3 \sigma_0 \sum_{n=0}^{\infty} A_n \left[ Q_{n-\frac{1}{2}}^2 (\cosh \sigma_0) \cdot \frac{1}{2} \frac{\cos n\psi}{(\cosh \sigma_0 - \cos \psi)^{5/2}} \right. \\
 &\quad \left. + \frac{dQ_{n-\frac{1}{2}}^2}{d(\cosh \sigma_0)} \frac{\cos n\psi}{(\cosh \sigma_0 - \cos \psi)^{3/2}} \right] \dots (5.5)
 \end{aligned}$$

Substituting (5.5) in (5.3) we get

$$\begin{aligned}
 I_b &= -\frac{1}{2} a \sinh^2 \sigma_0 \int_0^{2\pi} d\psi \sum_{n=0}^{\infty} A_n \left[ Q_{n-\frac{1}{2}}^2 (\cosh \sigma_0) \frac{1}{2} \frac{\cos n\psi}{(\cosh \sigma_0 - \cos \psi)^{3/2}} \right. \\
 &\quad \left. + \frac{dQ_{n-\frac{1}{2}}^2}{d(\cosh \sigma_0)} \frac{\cos n\psi}{(\cosh \sigma_0 - \cos \psi)^{\frac{1}{2}}} \right]
 \end{aligned}$$

Using the formula (Bateman 8, p. 464)

$$\int_0^{2\pi} \frac{\cos n\psi}{(\cosh \sigma - \cos \psi)^{\frac{1}{2}}} d\psi = 2\sqrt{2} Q_{n-\frac{1}{2}} (\cosh \sigma),$$

and that obtained from it by differentiation with respect to  $\cosh \sigma$

$$\int_0^{2\pi} \frac{\cos n\psi}{(\cosh \sigma - \cos \psi)^{3/2}} d\psi = 4\sqrt{2} \frac{dQ_{n-\frac{1}{2}} (\cosh \sigma)}{d(\cosh \sigma)}$$

we are led to the expression

$$I_b = -a\sqrt{2} \sinh^2 \sigma_0 \sum_{n=0}^{\infty} A_n \left[ Q_{n-\frac{1}{2}} \frac{dQ_{n-\frac{1}{2}}^2}{d(\cosh \sigma)} - Q_{n-\frac{1}{2}}^2 \frac{dQ_{n-\frac{1}{2}}}{d(\cosh \sigma)} \right]_{\sigma=\sigma_0} \dots (5.6)$$

Remembering that

$$\begin{aligned}
 A_0 &= \frac{1}{2\pi\sqrt{2}} \frac{Q_{-\frac{1}{2}} (\cosh \sigma_0)}{Q^2_{-\frac{1}{2}} (\cosh \sigma_0)} \\
 A_n &= \frac{1}{\pi\sqrt{2}} \frac{Q_{n-\frac{1}{2}} (\cosh \sigma_0)}{Q^2_{n-\frac{1}{2}} (\cosh \sigma_0)} \quad (\pi=1, 2, \dots)
 \end{aligned}$$

we see that (5.6) can be written

$$I_b = -\frac{a \sinh^2 \sigma_0}{\pi} (\frac{1}{2}B_0 + B_1 + B_2 + \dots), \dots \dots \dots (5.7)$$

where

$$B_n = \frac{Q_{n-\frac{1}{2}} (\cosh \sigma_0)}{Q_{n-\frac{1}{2}}^2 (\cosh \sigma_0)} \left[ Q_{n-\frac{1}{2}} (\cosh \sigma) \frac{dQ_{n-\frac{1}{2}}^2}{d(\cosh \sigma)} - Q_{n-\frac{1}{2}}^2 (\cosh \sigma) \frac{dQ_{n-\frac{1}{2}}}{d(\cosh \sigma)} \right]_{\sigma=\sigma_0}$$

$$= \frac{Q_{n-\frac{1}{2}} (\cosh \sigma_0)}{Q_{n-\frac{1}{2}}^2 (\cosh \sigma_0)} \text{Wr} [Q_{n-\frac{1}{2}} (\cosh \sigma), Q_{n-\frac{1}{2}}^2 (\cosh \sigma)]_{\sigma=\sigma_0}, \dots \quad (5.8)$$

where Wr stands for Wronskian.

Finally, from (5.3), (5.4), and (5.7)

$$\frac{Z}{aA} = \pi(1 - \coth \sigma_0) - \frac{\sinh^2 \sigma_0}{\pi} (\frac{1}{2}B_0 + B_1 + B_2 + \dots), \dots \quad (5.9)$$

where  $B_n$  is given by (5.8).

## VI. DISPLACEMENTS

Since  $\widehat{rr} = \widehat{zz} = \widehat{\theta\theta} = \widehat{rz} = 0$  identically, it follows from Hooke's law that the only non-vanishing strains are

$$e_{r\theta} = \frac{1}{2\mu} \widehat{r\theta}, \quad e_{z\theta} = \frac{1}{2\mu} \widehat{z\theta}, \dots \quad (6.1)$$

where  $\mu$  is the modulus of shear rigidity; both  $e_{r\theta}$  and  $e_{z\theta}$  are functions of  $r$  and  $z$  only.

Substituting from (6.1) into the displacement-strain relations in cylindrical polars (Sokolnikoff 9, p. 202) we obtain

$$\frac{\partial u_r}{\partial r} = 0 \dots \quad (6.2)$$

$$\frac{\partial u_\theta}{\partial \theta} + u_r = 0 \dots \quad (6.3)$$

$$\frac{\partial u_z}{\partial z} = 0 \dots \quad (6.4)$$

$$\frac{1}{r} \frac{\partial u_r}{\partial \theta} + \frac{\partial u_\theta}{\partial r} - \frac{u_\theta}{r} = \frac{\widehat{r\theta}}{\mu} \dots \quad (6.5)$$

$$\frac{\partial u_z}{\partial z} + \frac{\partial u_r}{\partial z} = 0 \dots \quad (6.6)$$

$$\frac{\partial u_\theta}{\partial z} + \frac{1}{r} \frac{\partial u_z}{\partial \theta} = \frac{\widehat{\theta z}}{\mu} \dots \quad (6.7)$$

Equations (6.2) to (6.7) are satisfied by the displacements

$$\left. \begin{aligned} u_r &= 0 \\ u_z &= c\theta \\ u_\theta &= \psi(r, z), \end{aligned} \right\} \dots \quad (6.8)$$

where the form of the function  $\psi(r, z)$  is determined by substituting (6.8) in (6.5) and (6.7); this gives, respectively,

and

$$\left. \begin{aligned} \frac{\partial \psi}{\partial r} - \frac{\psi}{r} &= \frac{1}{\mu} \widehat{r\theta} = \frac{A}{\mu} \frac{1}{r^2} \frac{\partial \varphi}{\partial z} \\ \frac{\partial \psi}{\partial z} + \frac{c}{r} &= \frac{1}{\mu} \widehat{\theta z} = -\frac{A}{\mu} \frac{1}{r^2} \frac{\partial \varphi}{\partial r} \end{aligned} \right\} \dots\dots\dots (6.9)$$

from (2.3)

that is

$$\left. \begin{aligned} r^2 \frac{\partial \psi}{\partial r} - r\psi &= \frac{A}{\mu} \frac{\partial \varphi}{\partial z} \\ -r^2 \frac{\partial \psi}{\partial z} - cr &= \frac{A}{\mu} \frac{\partial \varphi}{\partial r} \end{aligned} \right\} \dots\dots\dots (6.10)$$

Equating  $\partial/\partial r$  of the first to  $\partial/\partial z$  of the second of (6.10) shows that we must have

$$\frac{\partial^2 \psi}{\partial r^2} + \frac{\partial^2 \psi}{\partial z^2} + \frac{1}{r} \frac{\partial \psi}{\partial r} - \frac{1}{r^2} \psi = 0$$

that is

$$\left( \nabla^2 - \frac{1}{r^2} \right) \psi(r, z) = 0 \dots\dots\dots (6.11)$$

We are interested in the axial displacement  $u_z$  of any cross-section, which corresponds to the longitudinal contraction or extension of a helical spring. To determine it numerically in special cases, we require the value of the constant  $c$  in (6.8) and proceed to determine its relation to known parameters. This is done by finding expressions for the stresses from the displacements (6.8) and relating these to the stresses as formerly evaluated in equations (4.19) and (4.20).

Since displacements obey under a change of the coordinates system the same laws of transformation as the coordinates themselves, we move from  $u_r, u_z, u_\theta$  to  $u_\sigma, u_\psi, u_\theta$  by the equations

$$\left. \begin{aligned} u_\sigma &= \cos \gamma \, u_r + \sin \gamma \, u_z \\ &= \frac{h \partial \sigma}{\partial r} u_r + \frac{h \partial \sigma}{\partial z} u_z \\ u_\psi &= \sin \gamma \, u_r - \cos \gamma \, u_z \\ &= \frac{h \partial \sigma}{\partial z} u_r - \frac{h \partial \sigma}{\partial r} u_z \\ u_\theta &= u_\theta. \end{aligned} \right\} \dots\dots\dots (6.12)$$

Now, from (3.1) and (3.15)-(3.16)

$$\frac{\partial \sigma}{\partial r} \frac{1}{a} (1 - \cosh \sigma \cos \psi), \quad \frac{\partial \sigma}{\partial z} = \frac{1}{a} \sinh \sigma \sin \psi \dots\dots (6.13)$$

so that (6.8) and (6.12) give

$$\left. \begin{aligned} \frac{u_\sigma}{h} &= -\frac{c}{a} \theta \sinh \sigma \sin \psi \\ \frac{u_\psi}{h} &= \frac{c}{a} \theta (\cosh \sigma \cos \psi - 1) \end{aligned} \right\} \dots\dots\dots (6.14)$$



and solving (6.11) in the same way in which the homogeneous equation corresponding to (2.10) was treated we find that  $u_\sigma/h_3$  has, in toroidal coordinates, the form

$$\frac{u_\theta}{h_3} = \frac{(\cosh \sigma - \cos \psi)^{3/2}}{a \sinh \sigma} \sum_{n=0}^{\infty} Q_{n-\frac{1}{2}}^2 (\cosh \sigma) (a_n \cos n\psi + b_n \sin n\psi), \dots \quad (6.15)$$

where  $a_n, b_n$  are coefficients which will not be determined here.

In curvilinear coordinates  $\alpha_i$  strains are expressed in terms of displacements by the equations (Sokolnikoff(9), pp. 199-200):

$$e_{ii} = \frac{\partial}{\partial \alpha_i} \left( \frac{u_i}{h_i} \right) + \frac{1}{h_i} \left( \frac{\partial h_i}{\partial \alpha_1} \frac{u_1}{h_1} + \frac{\partial h_i}{\partial \alpha_2} \frac{u_2}{h_2} + \frac{\partial h_i}{\partial \alpha_3} \frac{u_3}{h_3} \right)$$

$$e_{ij} = \frac{1}{2h_i h_j} \left[ h_i^2 \frac{\partial}{\partial \alpha_j} \left( \frac{u_i}{h_i} \right) + h_j^2 \frac{\partial}{\partial \alpha_i} \left( \frac{u_j}{h_j} \right) \right] \text{ for } i \neq j.$$

Substituting for toroidal coordinates the expressions for  $h_1, h_2, h_3$  from (3.8) and  $u_\sigma, u_\psi$  from (6.14) we easily verify that

$$e_{\sigma\sigma} = e_{\psi\psi} = e_{\theta\theta} = e_{\sigma\psi} = 0 \text{ identically, as required.}$$

Further

$$2e_{\sigma\theta} = \frac{1}{h h_3} \left[ h^2 \frac{\partial}{\partial \theta} \left( \frac{u_\sigma}{h} \right) + h_3^2 \frac{\partial}{\partial \sigma} \left( \frac{u_\theta}{h_3} \right) \right]$$

$$= -\frac{c}{a} \sin \psi + \sinh \sigma \frac{\partial}{\partial \sigma} \left( \frac{u_\theta}{h_3} \right) \dots \dots \dots (6.16)$$

and

$$2e_{\psi\theta} = \frac{1}{h h_3} \left[ h^2 \frac{\partial}{\partial \theta} \left( \frac{u_\psi}{h} \right) + h_3^2 \frac{\partial}{\partial \psi} \left( \frac{u_\theta}{h_3} \right) \right]$$

$$= \frac{c}{a} \frac{\cosh \sigma \cos \psi - 1}{\sinh \sigma} + \sinh \sigma \frac{\partial}{\partial \psi} \left( \frac{u_\theta}{h_3} \right) \dots \dots \dots (6.17)$$

Now, from (4.19) it follows that

$$2e_{\sigma\theta} = \frac{\widehat{\sigma\theta}}{\mu} = -\frac{A}{2\mu a} \sin \psi + \frac{A}{a\mu} (\cosh \sigma - \cos \psi)^{3/2} \sum_{n=0}^{\infty} \left[ \frac{3 \sin \psi \cos n\psi}{2 (\cosh \sigma - \cos \psi)} \right.$$

$$\left. + (n \sin n\psi) A_n Q_{n-\frac{1}{2}}^2 (\cosh \sigma) \right] \dots \dots (6.18)$$

and from (4.20)

$$2e_{\psi\theta} = \frac{\widehat{\psi\theta}}{\mu} = \frac{A}{2\mu a} \frac{\cosh \sigma \cos \psi - 1}{\sinh \sigma} + \frac{A}{a\mu} \frac{(\cosh \sigma - \cos \psi)^{3/2}}{\sinh \sigma} \sum_{n=0}^{\infty}$$

$$\left[ \frac{\cosh \sigma (\cosh \sigma - 4 \cos \psi) + 3}{2 (\cosh \sigma - \cos \psi)} Q_{n-\frac{1}{2}}^2 (\cosh \sigma) + \sinh^2 \sigma \left( \frac{d Q_{n-\frac{1}{2}}^2 (\cosh \sigma)}{d (\cosh \sigma)} \right) A_n \cos n\psi \right]$$

$$\dots \dots \dots (6.19)$$

A comparison of (6.16) with (6.18) and of (6.17) with (6.19) gives the desired result:

$$c = A/2\mu \dots \dots \dots (6.20)$$

## VII. DISCUSSION OF RESULTS. CONCLUSIONS

The distribution of shearing stress along a horizontal diameter is shown for the particular case

$$R_0/\rho_0=4$$

in Figure 1 and around the circumference of this toro in Figure 2. The magnitude of the stresses was computed from (4.20); the series converge very rapidly and the first six terms ( $n=0$  to 5) gave an accuracy to the number of figures quoted in the tables. *Tables of the Associated Legendre Functions*(10) were employed for this part of the numerical work.

For the computation of the quantity

$$\text{Wr}[Q_{n-\frac{1}{2}}(\cosh \sigma_0), Q_{n-\frac{1}{2}}^2(\cosh \sigma_0)]$$

in expressions (5.8) and (5.9) for the resultant stress the Tables did not yield sufficient accuracy because of the loss of significant figures on subtraction. The expansion (Hobson(7), p. 304)

$$Q_{n-\frac{1}{2}}^m(x) = \frac{\Gamma(n+m+\frac{1}{2})\Gamma(\frac{1}{2})}{(-1)^m 2^{n+\frac{1}{2}} n! (x^2-1)^{\frac{1}{2}} x^{n-m+\frac{1}{2}}} F\left(\frac{n-m+\frac{3}{2}}{2}, \frac{n-m+\frac{1}{2}}{2}, n+1; \frac{1}{x^2}\right), \quad (7.1)$$

where

$$F(a, b, c; y) = \sum_{r=0}^{\infty} \frac{\Gamma(c)\Gamma(a+r)\Gamma(b+r)}{\Gamma(a)\Gamma(b)\Gamma(c+r)r!} y^r \quad (7.2)$$

was therefore used for numerical computation of  $\text{Wr}$ ;  $\Gamma(x)$  is the Gamma function of  $x$  and  $F(a; b; c; y)$  the hypergeometric function. The detailed manipulation will not be given here, but it is not difficult to verify the following results:

$$\text{Wr}(Q_{-\frac{1}{2}}(x), Q_{-\frac{1}{2}}^2(x)) = -\frac{1}{2}[\Gamma(\frac{1}{2})]^2 \Gamma\left(\frac{5}{2}\right) \left( \frac{2}{x^4} + \frac{4 \cdot 375}{x^6} + \frac{6 \cdot 9727}{x^8} + \frac{9 \cdot 7284}{x^{10}} + \frac{12 \cdot 6067}{x^{12}} + \dots \right)$$

$$\text{Wr}(Q_{\frac{1}{2}}(x), Q_{\frac{1}{2}}^2(x)) = -\frac{1}{8}[\Gamma(\frac{1}{2})]^2 \Gamma\left(\frac{3}{2}\right) \Gamma\left(\frac{7}{2}\right) \left( \frac{1}{x^6} + \frac{2 \cdot 625}{x^8} + \frac{4 \cdot 7065}{x^{10}} + \frac{8 \cdot 6125}{x^{12}} + \dots \right)$$

$$\text{Wr}(Q_{3/2}(x), Q_{3/2}^2(x)) = -\frac{1}{128}[\Gamma(\frac{1}{2})]^2 \Gamma\left(\frac{5}{2}\right) \Gamma\left(\frac{9}{2}\right) \left( \frac{2}{3} \frac{1}{x^8} + \frac{2 \cdot 0625}{x^{10}} + \frac{4 \cdot 1465}{x^{12}} + \dots \right)$$

$$\text{Wr}(Q_{5/2}(x), Q_{5/2}^2(x)) = -\frac{1}{4608}[\Gamma(\frac{1}{2})]^2 \Gamma\left(\frac{7}{2}\right) \Gamma\left(\frac{11}{2}\right) \left( \frac{1}{2} \frac{1}{x^{10}} - \frac{63}{5} \frac{1}{x^{12}} + \dots \right), \text{ etc.}$$

Values for the resultant stress are tabulated in Table 3 and the ratios of maximum and minimum stress to resultant stress in Table 4: Figure 3 embodies the results of the latter.

The stresses have been multiplied by appropriate factors before tabulation and graphing in order to show that our exact results tend asymptotically to the approximate results of engineering theory when the toro thickness is decreased

TABLE 3

AXIAL DISPLACEMENT OF ANY POINT IN THE TORE  
IN TERMS OF THE EXTERNAL FORCE  $Z$ ; CF.  
EQUATION (5.9)

$$\text{Axial Displacement } u_z = c\theta; \quad \frac{Z}{c} = \delta \cdot \frac{\pi\mu}{2} \frac{\rho_0^4}{R_0^3},$$

where  $\theta$  is Angular Coordinate of Any Point

$R_0/\rho_0$	$-\delta$
3	1.025
4	1.017
5	1.007
6	1.006
8	1.004
10	1.001

indefinitely. This theory gives the following formula for the axial lengthening or shortening of a helical spring of thin wire and small pitch with  $n$  turns (Geckeler 11, p. 192):

$$\Delta l = \frac{4nR_0^3}{\mu\rho_0^4} Z \dots\dots\dots (7.3)$$

clearly

$$n = \frac{\theta}{2\pi},$$

TABLE 4

MAXIMUM AND MINIMUM SHEAR STRESS FOR GIVEN EXTERNAL  
FORCE AND TORE DIMENSIONS; CF. EQUATIONS (4.20) AND  
(5.9) AND FIGURE 3

$$\text{Shear Stress } \psi\theta = K \cdot \frac{2R_0}{\pi\rho_0^3} Z$$

$\frac{R_0}{\rho_0}$	$K$	
	Maximum Shear Stress	Minimum Shear Stress
3	1.552	0.6571
4	1.376	0.7288
5	1.293	0.7800
6	1.237	0.8127
8	1.171	0.8554
10	1.136	0.8840

where  $\theta$  is the angular distance between the ends of the wire ; in our notation so that (7.3) becomes

$$u_z = \frac{2R_0^3}{\pi\rho_0^4} \theta \frac{Z}{\mu} \dots\dots\dots (7.4)$$

We have found that

$$u_z = c\theta \dots\dots\dots (6.8)$$

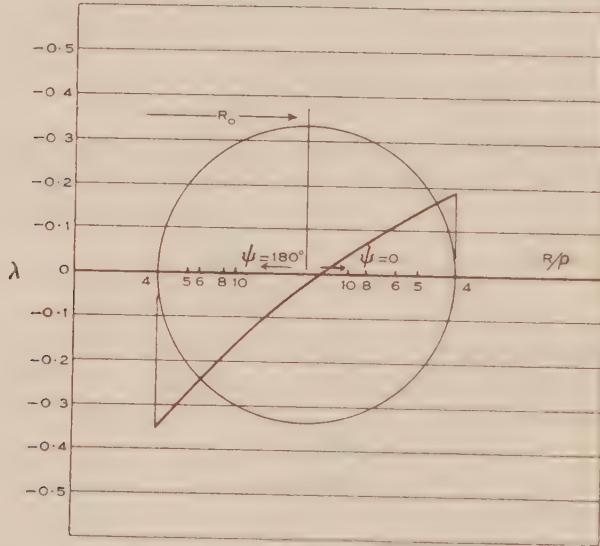


Fig. 3.—Distribution of shear stress  $\psi\theta = \lambda \frac{\mu c}{R}$  along a horizontal diameter of the torus  $\frac{R_0}{\rho_0} = 4$ .

Hence, identifying (7.4) and (6.8),

$$c = \frac{2R_0^3}{\pi\rho_0^4} \frac{Z}{\mu}$$

so that

$$\frac{2}{\pi\mu} \frac{R_0^3}{\rho_0^4} \frac{Z}{c} = 1 \dots\dots\dots (7.5)$$

Tabulating  $\frac{2}{\pi\mu} \frac{R_0^3}{\rho_0^4} \frac{Z}{c}$ , with our solution  $Z/\mu c$  from (5.9), as a function of  $R_0/\rho_0$  we should therefore expect this quantity to approach the value given by (7.5), i.e. unity, as the thickness of the torus diminishes, that is, as

$$\rho_0 \rightarrow 0 \text{ or } \frac{R_0}{\rho_0} \rightarrow \infty.$$

This is confirmed by Table 3 which also shows that even for thick torus

$$\frac{2}{\pi\mu} \frac{R_0^3}{\rho_0^4} \frac{Z}{c}$$

differs very little from unity.



The conclusion may be drawn that the engineering formula (7.3) is of wide applicability, giving good approximations for all cases likely to occur in practice.

The stress distribution around the circumference at any section will be nearly uniform for a thin tore, but not in general as Table 2 and Figure 2 show. We can verify, however, that our solution tends asymptotically to that for a right circular cylinder as the thickness of the tore tends to zero. The classical torsion theory gives for the torsional couple (Sokolnikoff 9, p. 120)

$$M = \mu \alpha I_0, \dots\dots\dots (7.6)$$

where  $I_0 = \frac{\pi \rho_0^4}{2}$  is the polar moment of inertia for a circular cross-section of a

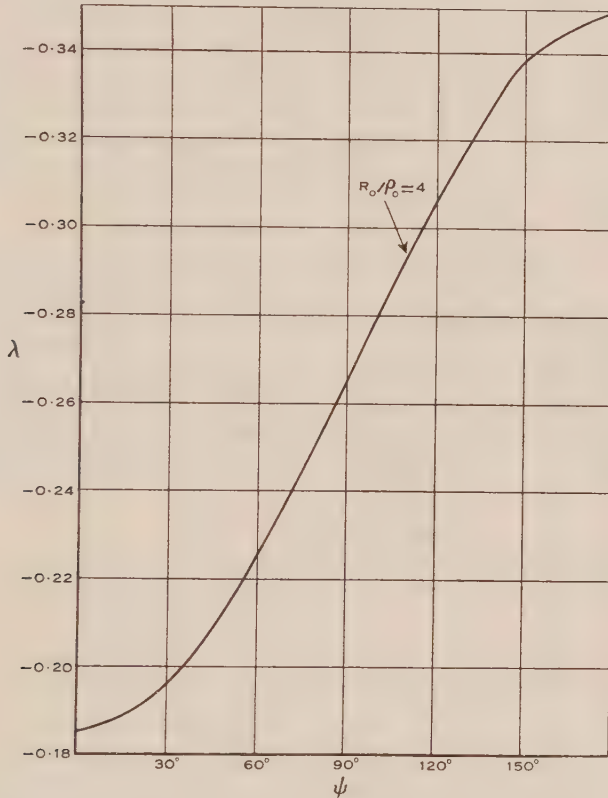


Fig. 4.—Distribution of circumferential shear stress  $\widehat{\psi\theta} = \lambda \frac{\mu c}{R_0}$   
for tore  $\frac{R_0}{\rho_0} = 4$

straight cylinder;  $\alpha$  is related to the tangential stress  $T$  as the boundary by (9, p. 121)

$$T = \mu \alpha \rho_0 \dots\dots\dots (7.7)$$

Clearly, in our notation,

$$M = ZR_0 \text{ and } T = \widehat{\psi\theta}$$

so that

$$\frac{\pi \rho_0^3}{2R_0} \frac{\widehat{\psi\theta}}{Z} = 1 \dots\dots\dots (7.8)$$

Dealing with (7.8) as we dealt with (7.5) above, i.e. tabulating and plotting

$$\frac{\pi \rho_0^3}{2R_0} \frac{\widehat{\psi\theta}}{Z}$$

as a function of  $R_0/\rho_0$ , we see from Table 4 and Figure 3 that

$$\left. \begin{aligned} \frac{\pi \rho_0^3}{2R_0} \frac{\widehat{\psi\theta}_{max}}{Z} \rightarrow 1 \text{ and } \frac{\pi \rho_0^3}{2R_0} \frac{\widehat{\psi\theta}_{min}}{Z} \rightarrow 1 \\ \text{with } R_0/\rho_0 \rightarrow \infty \end{aligned} \right\} \dots\dots\dots (7.9)$$

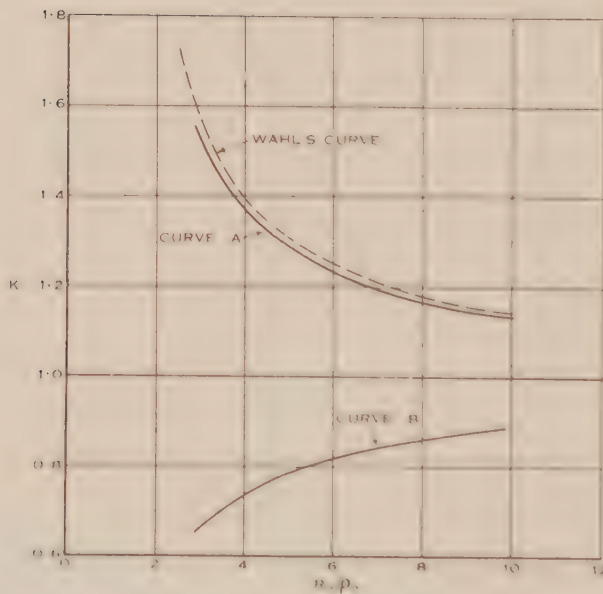
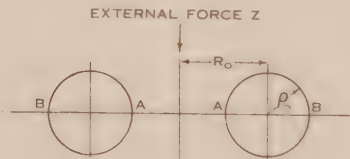


Fig. 5.—Maximum and minimum shear stress  $\widehat{\psi\theta} = K \frac{2R_0}{\pi \rho_0^3} Z$  for given external force  $Z$  and tore dimensions  $R_0, \rho_0$ . Curve  $A$  gives maximum stress (at points  $A$ ), curve  $B$  minimum stress (at points  $B$ ).

as required by the limiting case (7.8). Table 4 and Figure 3 show that for a given external force the maximum stress occurring in the tore considerably exceeds that given by ordinary torsion theory.

## VIII. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Aeronautics, C.S.I.R. It is desired to acknowledge the great benefit derived from discussions with Dr. R. C. T. Smith. Thanks are also due to Miss Claire Robbins for her assistance with the numerical work.

## IX. REFERENCES

- (1) SOUTHWELL, R. V.—Some practically important stress-systems in solids of revolution. *Proc. Roy. Soc. A* **180** : 367 (1942).
- (2) MICHELL, J. H.—The uniform torsion and flexure of incomplete tores, with application to helical springs. *Proc. Roy. Soc.* **31** : 130 (1899).
- (3) GÖHNER, O.—Schubspannungsverteilung im Querschnitt einer Schraubenfeder. *Ingen. Arch.* **1** : 619 (1930).
- (4) TIMOSHENKO, S.—“Theory of Elasticity.” (McGraw-Hill Book Co. : New York, 1934.)
- (5) WAHL, A. M.—Helical compression and tension springs. *J. Appl. Mech.* **2** : (1935).
- (6) WHITTAKER, E. T., and WATSON, G. N.—“A Course of Modern Analysis.” 4th Ed. (Cambridge Univ. Press : Cambridge, 1940.)
- (7) HOBSON, E. W.—“The Theory of Spherical and Ellipsoidal Harmonics.” (Cambridge Univ. Press : Cambridge, 1931.)
- (8) BATEMAN, H.—“Partial Differential Equations of Mathematical Physics.” (Dover Publications : New York, 1944.)
- (9) SOKOLNIKOFF, I. S.—“Mathematical Theory of Elasticity.” (McGraw-Hill Book Co. : New York, 1946.)
- (10) TABLES OF THE ASSOCIATED LEGENDRE FUNCTIONS. Mathematical Tables Project. (Columbia Univ. Press : New York, 1945.)
- (11) GECKELER, J. W.—Article on “Elastostatik” in Vol. VI (Mechanik der elastischen Körper) *Handb. der Physik.* (Julius Springer : Berlin, 1928.)

# THEORETICAL ASPECTS OF CLOUD DROP DISTRIBUTIONS

By E. B. KRAUS and BETTY SMITH\*

[*Manuscript received April 5, 1949*]

## *Summary*

A theoretical study indicates that the number and size of the drops formed in a cloud vary with the rate of cooling, the initial temperature, and the air pressure. The faster the cooling rate, the lower the initial temperature, and the lower the altitude, the greater is the number of drops and the smaller their size. The drop size spectrum also depends, to a large extent, on the number of available condensation nuclei. Furthermore, it tends to be widened by sedimentation and turbulence.

## I. INTRODUCTION

A number of investigators(1, 2, 3, 4, 5) have attempted to establish a relation between cloud forms and the number or the mean size of the droplets which they contain. The results vary considerably, perhaps because "cloud form" is an ambiguous concept for the purpose. Obviously there will be differences in the drop size distribution within one and the same cloud. In particular, sedimentation must play some role in stratus clouds and turbulence or convection patterns should cause local differences in cumulus clouds.

In spite of this ambiguity the various experimental results agree fairly well in showing qualitatively that cumuliform clouds are on the whole made up of a greater number of smaller drops than stratiform clouds. This can be related to the difference in the condensation rate. The size of the condensation nuclei is not uniform. The biggest or most effective nuclei will begin to form drops at a lower relative humidity than the others. If the cooling is slow enough, these first-formed drops should be able to take up all the further condensing moisture. They will grow and the relative humidity of the air will be kept down to the value which corresponds to their surface vapour pressure. That value is too low for the smaller nuclei to pass the threshold which is associated with the Kelvin curvature effect. The number of drops in slowly cooling air is thus restricted. On the other hand, when condensation occurs more rapidly, the moisture which becomes available cannot be diffused to the first formed drops as quickly. The relative humidity rises somewhat higher and some of the smaller nuclei can also grow into drops. The water which condenses later is distributed among a larger number of cloud particles, which therefore remain individually smaller.

A similar effect distinguishes the forms of condensation at initially different temperatures. More water becomes available when condensation occurs in a

\* Division of Radiophysics, C.S.I.R.O.



warm air mass than when a colder mass is cooled at the same rate. For example, the cooling by one degree of saturated air yields ten times as much liquid water at  $+20^{\circ}\text{C}$ . as at  $-10^{\circ}\text{C}$ . From the more rapid rate at which water becomes available in warm air, it might be expected that more nuclei would give rise to drops than in a colder air mass. The following computations suggest, however, the opposite effect. It appears that in warm air the first formed drops grow very rapidly, preventing the relative humidity from rising much above 100 per cent. even if the cooling rate is fast. In cold air, on the other hand, the first-formed drops cannot grow so rapidly. The relative humidity therefore rises somewhat higher, and more particles develop from the nuclear to the drop stage. If this is correct, a cold cloud should tend to contain more, but smaller, drops than a warm cloud.

It will be shown that height differences act in the same way as temperature differences. In a high cloud fewer drops are formed, but their average size should be bigger than in a low cloud of equal temperature and equal rate of cooling. Both the temperature and the height or pressure effect are slightly enhanced by the increase of the diffusion rate with the temperature and with the specific volume.

The treatment below is based on the assumption that drops take up instantaneously the temperature of the ambient air when the latter is saturated or very nearly saturated relative to a plane water surface. This same assumption is contained implicitly in the published calculations of the moist-adiabatic lapse rate. At the present state of observation we cannot say to what extent it approximates the actual state. Growing drops must be always slightly warmer than their environment and the difference between the dry and moist adiabatic lapse rate is fundamentally due to this fact. Boundary layer considerations seem to indicate, however, that the temperature difference tends to be very small for drops which fall with a velocity which can be represented by Stokes's law.

## II. THE GROWTH OF DROPS AS A FUNCTION OF THE TEMPERATURE AND THE RATE OF CONDENSATION

It is generally accepted by meteorologists that cloud drops form around hygroscopic, soluble nuclei. The mass of the dissolved material becomes small compared with that of the water, even before saturation is reached. The water vapour pressure at the drop surface is therefore given by the formulae of Kelvin and that of Raoult, which may be stated in the form

$$\ln E_r = \ln E_{\infty} + \frac{M}{RT\rho_1} \frac{2\sigma}{r} - \frac{c}{\rho}, \quad \dots\dots\dots (1)$$

where  $E_r$  = vapour pressure at drop surface,

$r$  = drop radius,

$E_{\infty}$  = saturation vapour pressure,

$M$  = 18 = molecular weight of water,

$R$  = universal gas constant,

$T$  = temperature,

$\rho_1$  = density of water,

$\rho$ =density of saturated water vapour,

$\sigma$ =surface tension,

$c$ =molecular ratio of concentration.

Assume that the solution droplet, which represents the nuclear stage of the drop, has a radius  $r_0$  when it is in equilibrium with exactly saturated air of vapour pressure  $E_\infty$ . We have then  $E_r=E_\infty$ , hence

$$c(r_0)=\frac{2M\sigma}{RT\rho_1}\frac{\rho}{r_0}.$$

When the drop grows the concentration decreases as  $r_0^3 r^3$ . Assuming a constant value of  $75 \text{ cm.}^3 \text{ sec.}^{-2}$  for  $\sigma/\rho_1$ , we obtain

$$c=c(r_0)\frac{r_0^3}{r^3}=\frac{2M}{RT}\frac{\sigma}{\rho_1}\frac{\rho r_0^2}{r^3}=3.25 \times 10^{-5} \frac{\rho r_0^2}{Tr^3}.$$

The values of  $E_r$  are very close to those of  $E_\infty$ . It is therefore possible to simplify as follows:

$$\ln E_r - \ln E_\infty = \ln \left( \frac{E_r - E_\infty}{E_\infty} + 1 \right) \approx \frac{E_r - E_\infty}{E_\infty} = h_r - 1,$$

where  $h_r$  is the relative humidity or the degree of supersaturation,  $E_r/E_\infty$ , at the drop surface.

Introducing the last two equations into (1) gives

$$h_r = 1 + \frac{3.25 \times 10^{-5}}{T} \frac{r^2 - r_0^2}{r^3} \dots \dots \dots (2)$$

Whilst developing from the nuclear state to full drop size each particle must pass through a stage at which the relative humidity at its surface is a maximum. The corresponding critical radius  $r_m$  can be computed by setting  $dh_r/dr=0$ . The result is

$$r_m = \sqrt{3} r_0 \dots \dots \dots (3)$$

and

$$h_{r_m} = 1 + \frac{2}{3} \frac{3.25 \times 10^{-5}}{Tr_m} = 1 + \frac{2.17 \times 10^{-5}}{Tr_m} \dots \dots \dots (4)$$

We define the first, or nuclear stage, in the growth of a cloud particle as being characterized by  $r < r_m$ , and the second or drop stage by  $r > r_m$ . It is reasonable to assume that in the nuclear stage the particle adjusts itself immediately, to all intents and purposes, in such a way as to make the vapour pressure at its surface equal to that of the surrounding air:

$$E_r = E \text{ or } h_r = h.$$

As the environment changes, the droplet will pass through a series of equilibrium stages, and its radius at any time will be given by equation (2). (This relation holds only for air that is relatively moist; for  $h < 0.7$  the droplet ceases to be a dilute solution and Raoult's law is not applicable.)

In the drop stage the volume increment becomes large enough to require a small but no longer negligible time interval for its formation. The change of volume is proportional to the radius, and not to the surface area as might be expected *a priori*. This fact has been shown experimentally and theoretically

by Houghton(6) and later in a different way by Findeisen(7). Using Findeisen's formula, which differs from Houghton's only by a numerical factor, we may write

$$d/dt\left(\frac{4}{3}\pi r^3\right)=24\cdot 3\rho_1 r\eta(q-q_r) \dots\dots\dots (5)$$

where  $\eta$ =coefficient of viscosity of air,

$q$ =specific humidity of the air,

$q_r$ =specific humidity at the drop surface.

Let  $q_\infty$  indicate the specific humidity at saturation vapour pressure. We have then

$$q=hq_\infty \qquad q_r=h_rq_\infty,$$

where  $h$  is the relative humidity of the cloud air. Introduction in (5) gives, after division by  $4\pi r^2$ ,

$$dr/dt=1\cdot 93\eta\rho_1 [(h-h_r)/r] q_\infty \dots\dots\dots (6)$$

The drops in a cloud will grow if more vapour molecules diffuse towards their surface than in the opposite direction. This can be the case only if  $q>q_r$ , and hence  $h>h_r$ . The difference may be very small, but it is nevertheless significant. The value of  $q$  at any time can be computed. It must be equal to  $q_0$ , the initial specific humidity of the air mass minus the amount of liquid water which has been formed per kg. of air.

$$q=q_0-\sum\frac{4}{3}\pi\rho(r^3-r_0^3)n\triangle r_0 \dots\dots\dots (7)$$

In this summation  $r$  is the radius of a particular drop the radius of which was  $r_0$  when saturation was first reached. The factor  $n\triangle r_0$  is the number of drops per kg. of air whose initial radius was between  $r_0-\frac{1}{2}\triangle r_0$  and  $r_0+\frac{1}{2}\triangle r_0$ , that is

$$r=r(r_0) \qquad n=n(r_0).$$

It is convenient to replace the summation by an integral. If furthermore we divide by  $q_\infty$  and assume the value unity for density of water  $\rho_1$ , we obtain

$$h=h^*-\frac{4\pi}{3q_\infty}\int_0^\infty (r^3-r_0^3)ndr_0, \dots\dots\dots (7')$$

where

$$h^*=\frac{q_0}{q_\infty},$$

which represents the degree of supersaturation the air would have at the same temperature had no liquid been deposited.

For cooling by only one or two degrees below the condensation temperature  $T_0$ , the value of  $q_0$  may be developed in a series of which only the first member needs to be considered :

$$q_0=q_\infty+\left[\frac{\partial q_\infty}{\partial T}\right]_0(T_0-T).$$

Elimination of  $q_0$  between the last two equations gives

$$h^*=1+\left[\frac{\partial(\ln q_\infty)}{\partial T}\right]_0(T_0-T)=1+\left[\frac{\partial(\ln E_\infty)}{\partial T}\right]_0(T_0-T),$$

where, in the last term, the change of air pressure during the cooling has been neglected.

With Magnus's formula(8)

$$\ln E_{\infty} = \frac{7.5(T-273)}{T-35.7} + 0.6609,$$

and hence

$$h^* = 1 + \frac{1779.8(T_0 - T)}{(T_0 - 35.7)^2} \dots\dots\dots (8)$$

Consider now a particular drop which had reached its critical radius at some as yet undetermined time  $t_m$ . From (6) it follows that at time  $t_1$  its radius will be

$$r = r_m + \int_{t_m}^{t_1} \frac{dr}{dt} dt = r_m + 1.93 \int_{t_m}^{t_1} \eta q_{\infty} \frac{h - h_r}{r} dt.$$

Introduction of the rate of temperature change  $\alpha = dT/dt$  gives

$$r = r_m + \frac{1.93}{\alpha} \int_{T_m}^{T_1} \eta q_{\infty} \frac{h - h_r}{r} dT, \dots\dots\dots (9)$$

where  $T_m$  is a function of  $C$  and  $r_0$  which can be computed and tabulated with the help of (4) and of tables for the saturation vapour pressure.

The quantities  $h$ ,  $h_r$ , and  $r_m$  can be eliminated with the help of equations (2), (3), (7'), and (8), giving

$$r = 1.73r_0 + \frac{1.93}{\alpha} \int_{T_m}^{T_1} \frac{\eta q_{\infty}}{r} \left[ 1779.8 \frac{T_0 - T}{(T_0 - 35.7)^2} - \frac{4.19}{q_{\infty}} \int_0^{\infty} (r^3 - r_0^3) n dr_0 - 3.25 \cdot 10^{-5} \frac{r^2 - r_0^2}{Tr^3} \right] dT \dots\dots (10)$$

For small pressure changes,  $\eta$  and  $q_{\infty}$  may be considered as depending on the temperature alone. The formula is therefore an integral equation with the variables  $T$  and  $r$ . If the rate of cooling  $\alpha$  is given, the value of  $r$  at a temperature  $T_1$  can be computed numerically step by step for any individual  $r_0$ .

### III. COMPUTED EXAMPLES OF CLOUD DROP DISTRIBUTIONS AND RELATIVE HUMIDITY CHANGES

In order to compute the size spectrum of clouds with the help of (10) it is necessary to know the number of nuclei of radius  $r_0$  in the air before condensation.

In the first instance it was assumed that the distribution of nuclei could be described by a probability curve for  $r_0$ . The curve was supposed to be symmetrical about  $r_0 = 5 \cdot 10^{-6}$  cm. and its height was determined by the requirement that the total number of nuclei should be

$$N = \int_0^{\infty} n dr_0 = 10^9 \text{ nuclei/kg. of air.}$$

This is equivalent to about 1000 nuclei per cc., which is the average found over the oceans and in free mountain air. The effect of different distributions will be discussed in Section V.



With this first distribution, nine cases were calculated. The rate of cooling was assumed to be  $\alpha=1^\circ\text{C./1800 sec.}$  corresponding to fog formation,  $\alpha=1^\circ\text{C./100 sec.}$  corresponding to rising air in fair weather cumuli, and  $\alpha=1^\circ\text{C./20 sec.}$  as in the up-currents of vigorous storm clouds. The initial temperature was taken to be  $-10^\circ\text{C.}$ ,  $0^\circ\text{C.}$ , and  $+15^\circ\text{C.}$

Figure 1 shows in detail the case  $\alpha=1^\circ\text{C./100 sec.}$  and  $T_0=0^\circ\text{C.}$  The diagram was calculated in steps of  $0.01^\circ\text{C.}$  The radius of each drop was assumed to be given by (2) as long as it had not reached the critical value  $r_m$ . Beyond this value its further growth was taken as being given by (10).

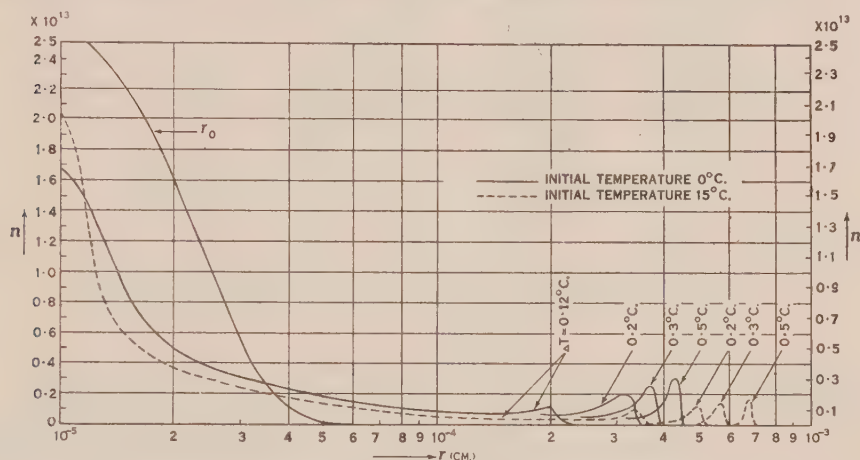


Fig. 1.—Distribution of drops in 1 kg. of air containing  $10^9$  condensation nuclei. Rate of cooling  $1^\circ\text{C./100 seconds.}$

Diagrams of this type are rather cumbersome and a more lucid representation may be obtained by graphs of the most frequent values of  $r$ , shown in Figures 2A, 2B, and 2C. The corresponding change of the relative humidity in the air is indicated by Figure 3.

It will be seen in Figure 3, for example, that in a fast and cold up-current the relative humidity should rise to more than 100.7 per cent. It would be difficult to corroborate this fact experimentally as it will last less than 20 seconds at the beginning of the condensation process. As soon as there is a sufficiently large liquid water surface available the relative humidity falls again.

The difference in the state of a cloud which may be produced by varying cooling rates is shown by the result that at a temperature of  $0^\circ\text{C.}$  a cooling of  $0.1^\circ\text{C.}$  at a rate of  $\alpha=1/1800$  will produce about  $5.5 \times 10^6$  drops per kg. of air with a most frequent radius of  $5.13 \times 10^{-4}$  cm. To produce the same most frequent drop size at a cooling rate  $\alpha=1/20$  would require a temperature drop of  $1.05^\circ\text{C.}$  A number of  $43.1 \times 10^6$  drops per kg. of air would be formed in this case.

Cooling by  $0.5^\circ\text{C.}$  at a rate  $\alpha=1/100$  produces  $18.7 \times 10^6$  drops of most frequent radius  $6.8 \times 10^{-3}$  cm. with an initial temperature of  $15^\circ\text{C.}$  and  $29.7 \times 10^6$  drops of most frequent radius  $4.2 \times 10^{-3}$  at  $0^\circ\text{C.}$  At  $-10^\circ\text{C.}$  the most frequent drop radius is only  $2.9 \times 10^{-3}$  cm.

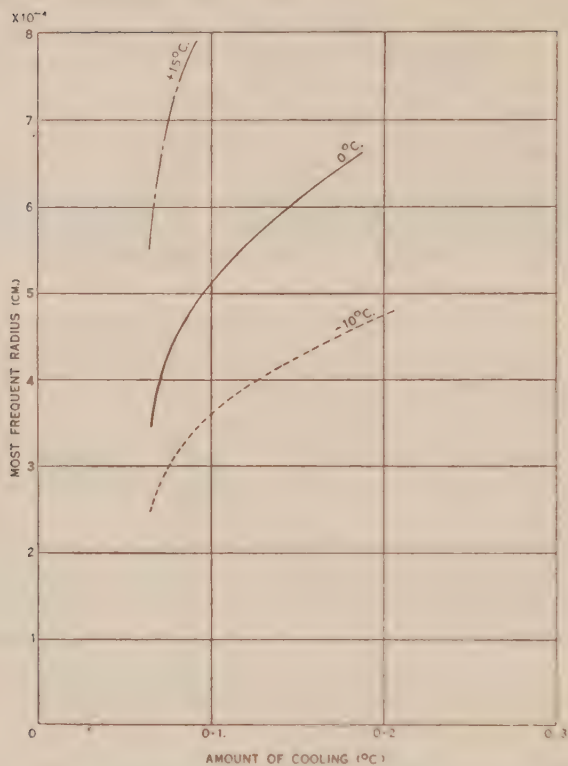


Fig. 2A.—Most frequent drop radius with cooling rate 1 °C./1800 seconds.

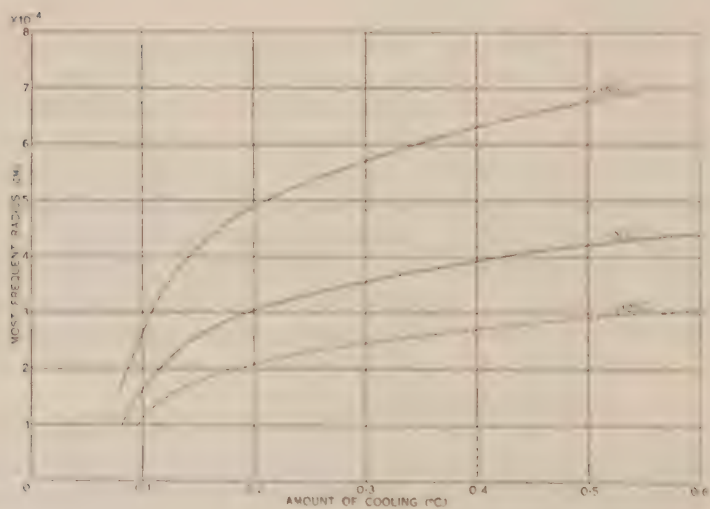


Fig. 2B.—Most frequent drop radius with cooling rate 1 °C./100 seconds.

The apparent paradox of a decrease in the number of drops with increasing moisture content of the air appears to be an interesting result which could explain, partly at least, the formation of rain in non-freezing clouds.

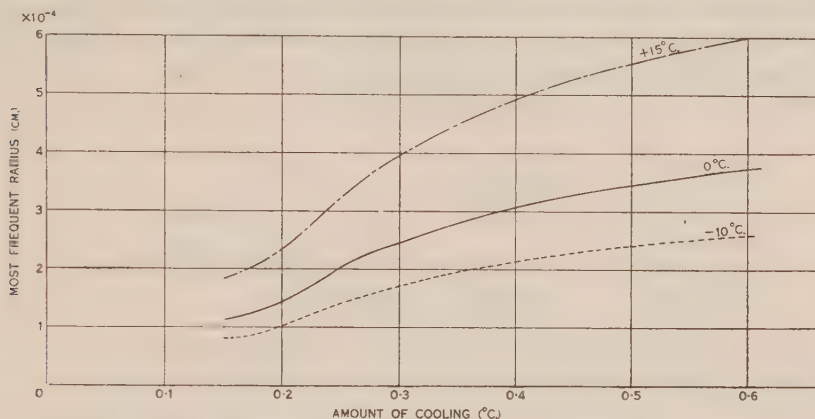


Fig. 2C.—Most frequent drop radius with cooling rate  $1^{\circ}\text{C}/20$  seconds.

#### IV. THE EFFECT OF THE CONDENSATION LEVEL UPON DROP SIZE

Equation (6) indicates that the rate of drop growth increases with the saturation specific humidity:

$$q_{\infty} = 0.622 E_{\infty} / P.$$

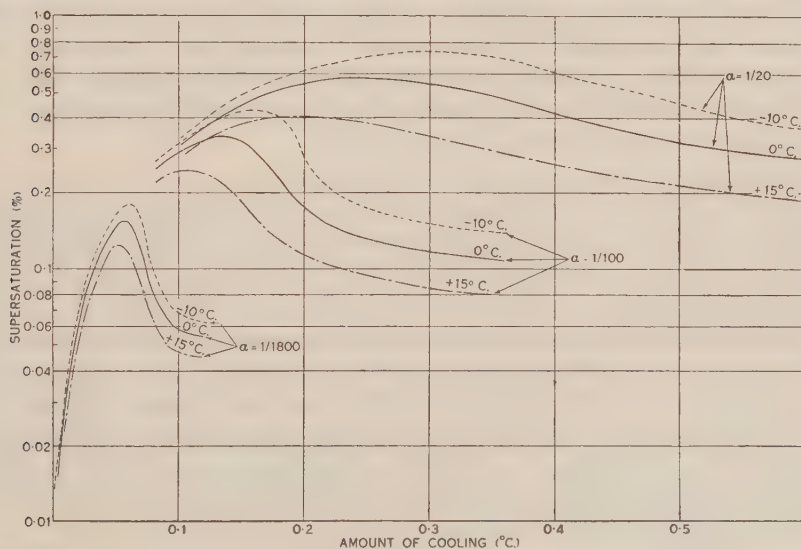


Fig. 3.—Supersaturation  $\frac{E - E_{\infty}}{E_{\infty}}$  as a function of the cooling rate  $\alpha$  and the initial temperature  $T_0$ .

Now the value of  $q_\infty$  is large in a warm air mass, and in such a case is associated with the rapid growth of drops. But  $q_\infty$  also varies with the air pressure, and in consequence at the same temperature has a higher value for lower pressures.

The dependence of drop growth on air pressure and hence on height is illustrated in Figure 4, which shows the most frequent radius plotted against temperature change for an initial temperature of  $0^\circ\text{C}.$ , and a cooling rate of  $1^\circ\text{C./100 sec.}$  The  $r_0$  distribution was assumed to be the same as in the previous cases, which referred to conditions at 700 mb.

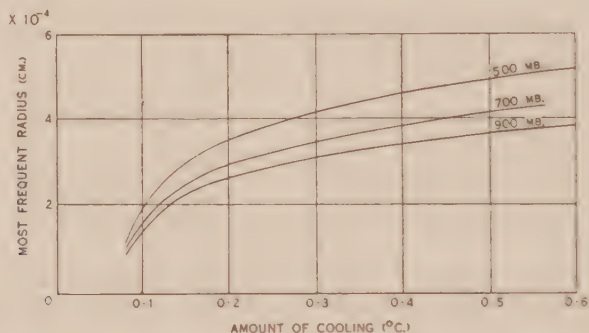


Fig. 4.—Most frequent drop radius at different pressure levels.

#### V. DIFFERENT DISTRIBUTION OF NUCLEI

The distribution of nuclei which formed the basis of the preceding calculations was based upon an estimate by Simpson(9), who indicates that atmospheric condensation nuclei vary in size from  $r_0$  less than  $5 \times 10^{-7}$  to more than  $5 \times 10^{-5}$ . Somewhat larger sizes were found to be prevalent in a series of observations by Dessens(10). The above computations were therefore repeated with the assumption of a ten times larger radius for all the nuclei. The results show that after a short time the size and the number of cloud droplets will be the same whatever the initial size of the nuclei.

Of much greater importance are variations in the number of nuclei. Dessens indicates that nuclei multiply when air dries out by crystallization and splitting of solution droplets. This process will be most active in anticyclones above the surface layers and over continents in general. On the other hand, the number of nuclei in surface air over the ocean should tend to remain small. This is confirmed by observations which show values as low as two per cc. over the seas against a minimum of 180 per cc. inland(11). Figure 5 shows the distribution of cloud drops in air which contained  $10^7$  nuclei per kg. or about 10 per cc. The proportion of the various sizes of nuclei and the conditions of condensation are the same as those of Figure 1. It will be seen immediately that the resulting distribution of cloud drops is much narrower and the most frequent radius considerably larger than in the former case.

The great difference which must be associated with a varying number of nuclei can be illustrated more strikingly and more logically by comparing the volumes of drops or their numbers, rather than the radii. Table 1 gives the



number of drops which should participate in the condensation process at different rates of cooling from an initial temperature of 0 °C.

Reduction in the number of nuclei has the greatest effect in the case of rapid condensation where the majority of nuclei may be used up. It is relatively unimportant in fog formation, which affects only the biggest nuclei in either case.

TABLE 1

$N$	1 °C./1800 sec.	1 °C./100 sec.	1 °C./20 sec.
$10^9$	$5.5 \times 10^6$	$29.7 \times 10^6$	$43.1 \times 10^6$
$10^7$	$1.9 \times 10^6$	$5.2 \times 10^6$	$6.0 \times 10^6$

The variation of the relative humidity is shown in Figure 6. The supersaturation could easily be more than four per cent. in warm air which is poor in nuclei and is rising rapidly.

## VI. THE EFFECT OF SEDIMENTATION AND TURBULENCE

Whilst all the drops are small enough to be carried upwards with a rising current, sedimentation tends to increase the difference between the radius of the largest and the smallest drops.

It is easily seen that a small drop will be carried up almost with the speed of the ascending current, whilst a larger one will take longer to rise the same distance and therefore has more time for its growth.

The change of cloud drop radius is represented by the equation

$$\frac{rdr}{dt} = \frac{24.3}{4\pi} \rho_1 \eta (q - q_r) \dots \dots \dots (5')$$

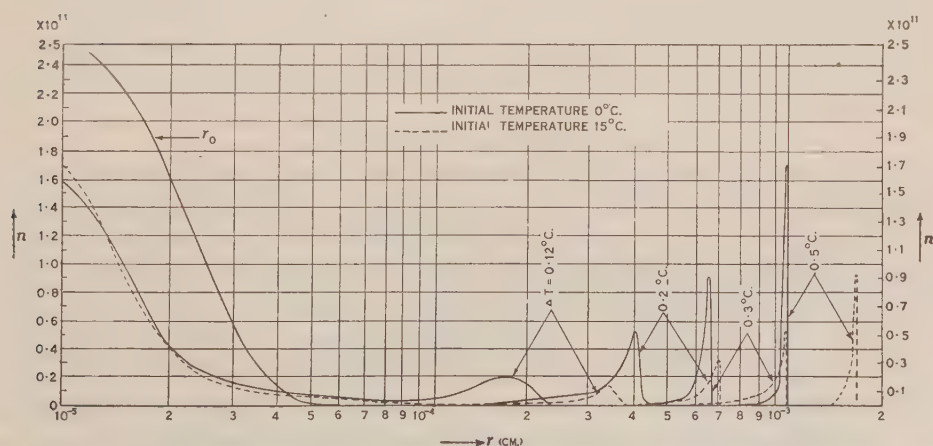


Fig. 5.—Distribution of drops in 1 kg. of air containing  $10^7$  condensation nuclei. Rate of cooling 1 °C./100 seconds.

By Stokes's law, its vertical displacement  $L$  in an up-current of constant velocity  $W$  is given by

$$\frac{dL}{dt} = W - \frac{2\rho_1 g r^2}{9\eta} \dots\dots\dots (11)$$

These two equations could be evaluated numerically in the same way as before. To investigate the effect of sedimentation alone it is, however, sufficient and much simpler to assume a constant mean value  $\overline{\Delta q}$  for  $q - q_r$ . With this assumption (5') can be integrated immediately, giving

$$\frac{1}{2}(r^2 - r_1^2) = \frac{24.3}{4\pi} \rho_1 \eta \overline{\Delta q} t, \dots\dots\dots (12)$$

where  $r_1$  is an integration constant.

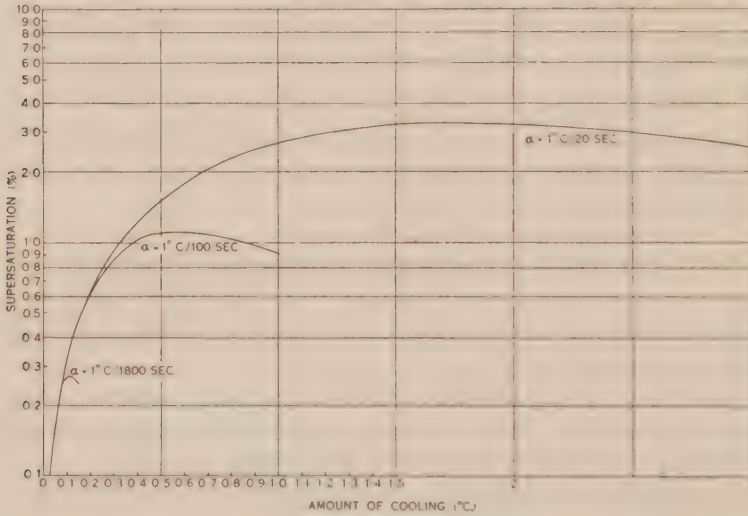


Fig. 6.—Supersaturation in air containing  $10^7$  nuclei/kg.

We combine (11) and (12), setting

$$\frac{2\rho_1 g}{9\eta} = \beta \text{ and } \frac{24.3\rho_1^2 g}{18\pi} \overline{\Delta q} = \gamma, \dots\dots\dots (13)$$

with the result

$$dL/dt = W - \beta r_1^2 - \gamma t.$$

Let  $r_1$  now indicate the drop radius at a level  $z_1$  and  $r$  its radius at a level  $z$ . Integration gives

$$z - z_1 = (W - \beta r_1^2)t - \gamma t^2.$$

The time can be eliminated from this last equation with the aid of (12), and the resulting equation, when solved for  $r$ , gives

$$r^2 = \frac{1}{\beta} [W \pm \sqrt{(W - \beta r_1^2)^2 - 4\gamma(z - z_1)}] \dots\dots\dots (14)$$

for the radius of a drop at a level  $z$ , when it had the radius  $r_1$  at the level  $z_1$ . The minus sign gives the case of a drop which rose straight from  $z_1$  to  $z$ ; the plus sign that of a drop which rose through  $z$  to some higher level until it became so large as to fall back to  $z$ .

If it is now assumed that the drop distribution in a small volume of air at  $z_1$  is constant, i.e. if the up-current arrives from below always with the same number of drops of equal size, the distribution at any other level can be computed from (14). The result of such a computation is shown in Figure 7 for a height difference  $z - z_1 = 10^5$  cm. and  $w = 50$  cm./sec. The broadening of the distribution of drops as compared with the case of no sedimentation, in which all drops rise with the speed of the current, is clearly indicated by the diagram. The two curves farthest to the right represent the set of falling drops.

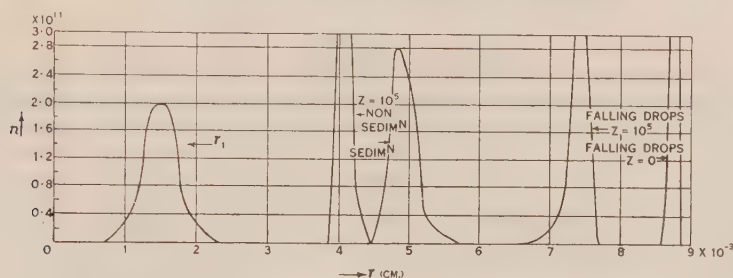


Fig. 7.—The effect of sedimentation on drop size distribution.

An air current which rises through a cloud 3000 ft. in thickness with a base at 700 mb. and  $0^\circ\text{C}$ . should deposit about 2 g. of liquid water per kg. of air. In reality we do not find such a high content of liquid water in clouds. Mixing with adjacent air must, therefore, be an important factor. In the present theory this effect has been neglected; hence the computations represent only qualitative tendencies and not real conditions.

Turbulent mixing must tend to broaden the drop size spectrum by mixing together drops of differing history and hence of different size. The degree of mixing is difficult to estimate. It must depend largely on the minimum size of the turbulent eddies. The more vigorous the turbulence, the smaller are the smallest eddies(12).

In particular, mixing with surrounding drier air will tend to produce drops with a relatively small radius at the periphery of an up-current.

## VII. CONCLUSION

Clouds have frequently been called "atmospheric colloids". Their state can be described completely only by a representation of the fields of motion, temperature, pressure, and humidity as well as by the distribution and the sizes of suspended drops and nuclei. These variables are not independent of each other, as the fields of motion, for example, may be determined by temperature and pressure through the hydrodynamic equations. The present study deals with some of the relations that affect the drop size spectrum and it may be summarized as follows.

- (i) The mode or most frequent radius of drops is larger, the slower the rate of cooling, the higher the initial temperature, and the lower the pressure.

- (ii) The number of initially available nuclei can greatly affect the state of a cloud. The number of nuclei appears to be smallest in the air over the oceans. Together with the conditions of the preceding paragraph, this might have a bearing on the formation of rain from non-freezing clouds.
- (iii) Sedimentation and mixing tend to broaden the drop size spectrum. Within an up-current and in the absence of falling drops the former tends to accentuate the growth of the largest drops, whilst the latter tends to increase the number of small drops.

#### VIII. ACKNOWLEDGMENT

The work described in this paper was carried out as part of the research programme of the Division of Radiophysics, C.S.I.R.O.

#### IX. REFERENCES

- (1) BRICARD, T.—*Météorologie* **1939**: 83 (1939).
- (2) MAZUR, J.—Air Minist. Met. Res. Cttee. MRP No. 109 (1943).
- (3) LANGMUIR, I.—Supercooled water droplets in rising currents of cold saturated air. Gen. Elect. Res. Lab. P2984 (6), Precipitation Static Studies, Oct. 1943-Aug. 1944.
- (4) DIEM, M.—*Ann. Hydrogr.* **70**: 142 (1942).
- (5) WEICKMANN, H.—*Beitr. Phys. frei. Atmos.* **28** (1): 12 (1945).
- (6) HOUGHTON, H. G.—*Physics* **4**: 419 (1933).
- (7) FINDEISEN, W.—*Met. Z.* **56**: 453 (1939).
- (8) ALBRECHT, F., ET AL.—Handb. Meteorologischen Instrumente. p. 201. (Edited by E. Kleinschmidt.) (Springer: Berlin, 1935.) (Lithoprint, Edwards: Ann Arbor, 1944.)
- (9) SIMPSON, G.—*Quart. J. Roy. Met. Soc.* **67**: 99 (1941).
- (10) DESSENS, H.—*Ann. Géophys.* **2**: 276 (1946).
- (11) LANDSBERG, H.—*Beitr. Geophys.* **3** (suppl.): 155 (1938).
- (12) BATCHELOR, G. K.—*Proc. Camb. Phil. Soc.* **43** (4): 533-59 (1947).



# THE CONTRACTION OF NUCLEAR EMULSIONS

By S. L. MARTIN\*

[Manuscript received June 21, 1949]

## Summary

A technique for measuring the fractional change in thickness which occurs when a nuclear emulsion is processed, is described. The method involves the registration in the emulsion of the tracks of  $\alpha$ -particles from a radiothorium source and this permanent record can be used to follow any subsequent thickness variations. Some observations on Ilford C2 emulsions, 50 $\mu$  in thickness, have been made and these show considerable variations in thickness over a period of time. Tests indicated that some of the change is probably due to humidity variations. The application of this method to the monitoring of variations in  $k$  for emulsions, in which deuterons or other tracks have been registered, is described.

## I. INTRODUCTION

It is known† that nuclear research emulsions when processed and dried shrink to about half their original thickness. During the progress of research on  $n$ - $d$  scattering in which Ilford C2 nuclear research emulsions were being used, the question arose as to whether the shrinkage ratio,  $k$ , given by

$$k = \frac{\text{thickness after processing}}{\text{thickness before processing}}$$

remained constant or was affected by changes in atmospheric conditions. Since preliminary tests using a micrometer dial gauge indicated that  $k$  varied from part to part of the same plate and between plates from the same batch, it was decided to investigate the matter using an adaptation of the method suggested by Powell *et al.* Suppose an  $\alpha$ -particle enters the emulsion along a direction  $AB$  (Fig. 1) making an angle  $\beta$  with the surface; when the plate has been developed, fixed, and dried, shrinkage reduces the inclination of the track to an angle  $\gamma$  given by

$$DC/DB = \tan \gamma / \tan \beta.$$

## II. EXPERIMENTAL

The apparatus employed (Fig. 2) consisted of a closed cylindrical brass container whose two parts,  $M$  and  $N$ , were held together by wing nuts and bolts, the joint between them being sealed by means of a rubber ring,  $R$ . The nuclear plate,  $P$ , inclined at 63° to the axis, was exposed to the  $\alpha$ -particles from a radiothorium source,  $S$ , of strength about 10 microcuries deposited on a platinum

\* Physics Department, Melbourne Technical College.

† POWELL, C. F., OCCHIALINI, G. P. S., LIVESEY, D. L., and CHILTON, L. V.—*J. Sci. Instrum.* 23: 102 (1946).

disk of 2 mm. diameter. During an exposure the pressure inside the container was reduced to approximately 1 mm. of mercury, the actual pressure at any instant being obtained from a dial vacuum gauge, *G*, of the Bourdon type. After processing according to the directions given by the manufacturer the plates were examined with a Cooke, Troughton, and Simms's microscope, type M4001, using an oil-immersion objective and an overall magnifying power of 2000. This microscope is fitted with a fine-focusing adjustment, the divided head of



Fig. 1.

which allowed the depth *DC* (Fig. 1) of a point on a track to be found. The projected track length *AD* was found by means of a graticule scale in the eyepiece.  $\tan \beta$  is given by (see Fig. 3)

$$\tan \beta = \frac{SN}{PN} = \frac{OS \sin \theta}{\sqrt{(OS \cos \theta - x)^2 + y^2}},$$

where *x* and *y* (measured by stage micrometers) are the coordinates relative to the plate centre, *O*, of the track being measured. In the apparatus used  $\theta = 63^\circ$

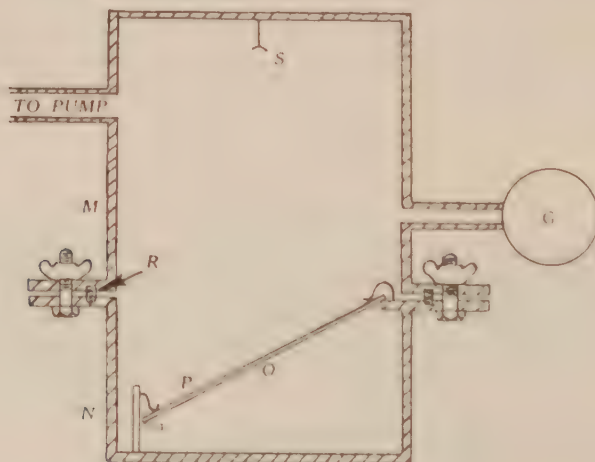


Fig. 2.

and  $OS=16.05$  cm. With the source used, a 15 minutes exposure of the plate to the  $\alpha$ -particles gave two or three tracks of average projected length about  $15\mu$  in each field of view.

### III. RESULTS

Tests carried out on several plates showed that the shrinkage ratio was not constant from day to day. Table 1 shows the variations of  $k$  which occurred for the central portions of four plates,  $A$ ,  $B$ ,  $C$ , and  $D$ , all of nominal thickness  $50\mu$ .

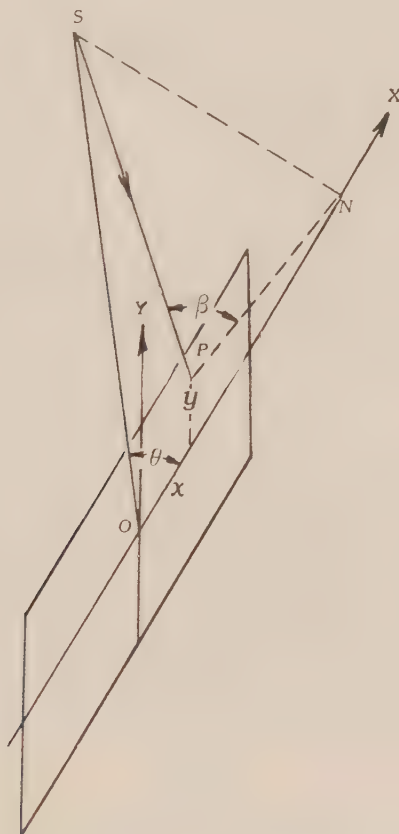


Fig. 3.

It will be seen that there is a considerable variation. Plate  $A$ , although from the same batch as  $B$ , has, generally, a higher  $k$  value. That the variation cannot be completely due, if at all, to possible variations in the developing and fixing procedures, is shown by the fact that  $k$  varies from part to part of the same plate. For instance, measurements made at approximately the same time at three different regions of plate  $A$  gave  $0.45$ ,  $0.54$ , and  $0.61$  respectively. To check whether the changes in  $k$  from plate to plate could be due, at least in part, to variations in the humidity of the air, some of the plates were kept for

TABLE 1

Plate A		Plate B		Plate C		Plate D	
$k$	No. of Days after Drying	$k$	No. of Days after Drying	$k$	No. of Days after Drying	$k$	No. of Days after Drying
$0.54 \pm 0.02$	1	$0.46 \pm 0.02$	4	$0.53 \pm 0.02$	1	$0.42 \pm 0.02$	32
$0.52 \pm 0.02$	6	$0.49 \pm 0.02$	7	$0.49 \pm 0.02$	30	$0.42 \pm 0.02$	49
$0.61 \pm 0.02$	15	$0.53 \pm 0.02$	14			$0.46 \pm 0.02$	62
$0.59 \pm 0.02$	16	$0.48 \pm 0.02$	21			$0.52 \pm 0.02$	81
$0.60 \pm 0.02$	18	$0.50 \pm 0.02$	50			$0.45 \pm 0.02$	82

periods ranging from 3 to 24 hours in an atmosphere saturated with water vapour. In every case  $k$  was found to have increased to about 1.3. After these plates were measured they were then placed in a dry atmosphere for about 24 hours and, as shown in Table 2, each one returned to approximately its original  $k$  value.

#### IV. ACCURACY OF MEASUREMENTS

The accuracy of measurement of  $k$  is mainly limited by the grain size which, in the case of the Ilford C2 nuclear emulsion, is of the order of  $0.6\mu$ ; no matter where the ionizing particle passes through the grain, the whole grain becomes developable. Because of this uncertainty as to the actual part of the grain on which to focus, together with that due to the depth of focus of the microscope it was considered that the depth measurement was known only to within about 7 per cent. Depth measurements were made only over the straight portions of tracks which themselves were substantially straight. The projected length of the track could be determined to within about 1 per cent. In view of the

TABLE 2  
VALUES OF  $k$

Plate	Before Placing in Saturated Atmosphere	Immediately after Removal from Saturated Atmosphere	After being First in Saturated Air and then in Dry Air
$E$	0.58	1.39	0.60
$F$	0.50	1.31	0.49



fairly large possible error in the depth measurement the value quoted for  $k$  at any point is the average calculated from observations on at least 10 tracks occurring in a small region of the part of the plate under observation.

#### V. APPLICATION OF THE METHOD TO MONITORING PLATES

This method may be used for monitoring day-to-day variations in  $k$  on plates in which other tracks (e.g. deuteron tracks) have been registered. In this case it is sometimes difficult to distinguish whether a track is due to an  $\alpha$ -particle or to another ionizing particle. The source should be placed outside the container since the radioactive gas, thoron, is one of the disintegration products of radiothorium and this would produce a number of tracks not coming directly from the source. In an apparatus used by the author the  $\alpha$ -particles entered through a mica window of stopping power 0.8 cm. and 3 mm. in diameter, the mounting of the source being such that light could not enter through the mica window and fog the plate. Even with the source external there were still a number of tracks not having the correct direction and it was assumed that these were due to particles scattered from the walls of the container. On this account when the method was used to mark plates carrying deuteron tracks whose angular distribution was being found, a metal mask having a few holes each about 1 cm. square was placed over the nuclear plate while it was exposed to the  $\alpha$ -particles. This ensured that the marking tracks occurred only over a few known regions; measurements on the deuteron tracks were made only on the parts covered by the metal mask and close to the region for which the  $k$  value had been found.

#### VI. ACKNOWLEDGMENTS

This work was carried out at University College, London, and the author wishes to thank Professor Massey for his interest in the investigation, Dr. E. H. S. Burhop for helpful discussions, and Mr. J. Tuckfield for help in assembling the apparatus. It is a pleasure to acknowledge the financial help received from the Council of Melbourne Technical College, from which institution the author was on leave of absence.

# THE SECOND VISCOSITY COEFFICIENT IN RHEOLOGICAL SYSTEMS

By R. C. L. BOSWORTH\*

[Manuscript received May 16, 1949]

## Summary

Viscosity in liquids may be regarded as that property which corresponds to elasticity in solid systems, and the coefficient of viscosity is expected to show as many independent components as the modulus of elasticity. For isotropic systems two components are expected. These two components are denoted by  $\eta$  and  $\zeta$ . The Stokes relationship between the two coefficients, viz.

$$\zeta = -\frac{2}{3}\eta$$

is compared with the Cauchy-Poisson relationship between the two moduli of elasticity of an isotropic solid, and is claimed as of no more general application.

Indeed, experience in sound absorption leads to the conclusion that

$$\zeta \gg \eta,$$

and many measurements on the viscosity of plastic materials appear tacitly to assume that

$$\zeta = \infty.$$

$\zeta$ , however, has a finite value for systems which undergo a progressive change in density at constant pressure. For such systems it is shown that the coefficient of viscosity, as determined by a tensile testing method, while always less than that as determined by torsion testing methods is greater than two-thirds of the latter value. The bulk viscosity, measured by the bulk stress required to produce a progressive density change, is shown to be related to the difference between the shear (or torsional) and the tensile (or flexural) viscosities.

Measurements of all three viscosities carried out on a specimen of solid carbon dioxide gives a satisfactory agreement with the theory of small deformations. The values recorded are:

Coefficient of viscous traction	$10.5 \times 10^{10}$ poises,
Torsional viscosity	$4.2 \times 10^{10}$ poises,
Bulk viscosity	$5.4 \times 10^{10}$ poises (measured),
or	$7.0 \times 10^{10}$ poises (computed).

Both flexural and torsional flow curves are of the Bingham type. Corrected yield points are

$6.9 \times 10^6$  baryes for tensile stress  
and  $4.8 \times 10^6$  baryes for shear stress.

## I. INTRODUCTION

An important section of the science of rheology is concerned with the behaviour of those types of matter which simultaneously show the properties of elasticity and of viscosity. Elasticity is usually regarded as a characteristic

\* Colonial Sugar Refining Co., Sydney.

property of solids, viscosity as characteristic of fluids. The fact that these properties are closely related follows from the definitions:

Modulus of elasticity = stress/strain,

Coefficient of viscosity = stress/rate of change of strain.

Dimensionally, therefore, the two properties differ only by a single power of time and we should expect coefficient of viscosity to show the same spatial behaviour as modulus of elasticity. Whereas, however, an arbitrary body is generally recognized as possessing several independent moduli of elasticity an arbitrary fluid is seldom, in modern treatments of the subject of viscosity, regarded as possessing more than one coefficient of viscosity. While there are, as will be seen below, historical reasons for this simplified approach to the subject, it is to be expected that such a treatment will lead to inconsistencies when viscosities are measured by methods involving different stress components. Methods of measuring the viscosity of highly viscous rheological systems frequently do employ stress components which differ from method to method, and it would appear desirable to review this subject theoretically and by comparing the viscosities of a system which can be examined under different stress components.

Elasticity is a symmetrical tensor quantity of the fourth rank. An anisotropic solid body, such as a triclinic crystal, has therefore 21 independent moduli of elasticity(1). With increasing symmetry a solid body has fewer and fewer independent moduli. Thus a tetragonal body has six, a cubic solid three, and an isotropic (or amorphous) solid two only. The elastic properties of isotropic solids are usually specified in terms of two of the following four elastic quantities: Young's modulus ( $Y$ ), bulk modulus ( $K$ ), shear modulus or rigidity ( $G$ ), and the Poisson ratio ( $\sigma$ ). These four quantities are connected by the two equations

$$\frac{1}{Y} = \frac{1}{3G} + \frac{1}{9K}, \dots \dots \dots (1)$$

and

$$\sigma = \frac{3K - 2G}{6K + 2G} \dots \dots \dots (2)$$

Any other independent equation between a pair of these quantities would reduce the elasticity of an isotropic solid to the rank of a scalar. One such attempted relationship is given by the Cauchy-Poisson equation, viz.

$$G = \frac{3}{5}K \dots \dots \dots (3)$$

This equation was derived on the assumption that the forces between the molecules of the solid are purely central forces. On combining equations (2) and (3) we get

$$\sigma = 0.25.$$

Experimental values of  $\sigma$  vary from 0.08 (for concrete) to about 0.49 (for rubber) and only a very limited class gives values in accordance with the Cauchy-Poisson relationship.

## II. THE STOKES. VISCOSITY RELATIONSHIP

Turning now to the case of fluids we must expect, since the formation of a time derivative does not alter the tensor nature of a quantity, that the viscosity



of an anisotropic fluid would exhibit 21 independent components. While it is unlikely that the experimental realization of all these properties would be possible in the present state of experimental technique there is no reason why the two independent components of the isotropic (or ordinary amorphous) fluid should not be known for most of the systems which have been studied. The expression for the stress tensor components in a fluid subject to varying velocities was given by Stokes who in 1845 first recognized the existence of two possible coefficients of viscosity. The stress component  $p_{ij}$ , or the  $x_i$  momentum transferred in unit time across unit plane normal to the  $x_j$  direction is given by

$$-p_{ij} = -p_0 \delta_{ij} + \eta \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right) + \zeta \delta_{ij} \operatorname{div} u, \dots\dots\dots (4)$$

where  $u_i, u_j$  are the velocity components and  $x_i, x_j$  the coordinates.  $\delta_{ij}$  is the Kronecker symbol

$$\begin{aligned} \delta_{ij} &= 1 & j &= i \\ \delta_{ij} &= 0 & j &\neq i, \end{aligned}$$

$p_0$  is the equilibrium pressure which is a function of the density but not of the time derivative of the density,  $\eta$  and  $\zeta$  are the two viscosity coefficients.

As will be seen from equation (3) the second coefficient ( $\zeta$ ) contributes only to the principal components of the stress tensor. The sum of the principal components  $p_{11} + p_{22} + p_{33}$  or  $3p$  is

$$\begin{aligned} 3(p - p_0) &= (2\eta + 3\zeta) \operatorname{div} u, \\ \text{or} \quad p - p_0 &= \left( \frac{2}{3}\eta + \zeta \right) \operatorname{div} u, \\ &= \kappa \operatorname{div} u, \dots\dots\dots (5) \end{aligned}$$

where  $\kappa$  is written for  $\frac{2}{3}\eta + \zeta$ . Stokes(2), who incidentally was thinking in terms of gases and certainly not in terms of rheological systems, argued that internal motion of the molecules could not produce frictional forces which would influence the external pressure. Consequently,

$$p = p_0, \quad \kappa = 0,$$

giving us a relationship between  $\eta$  and  $\zeta$ , viz.

$$\eta = -\frac{2}{3}\zeta \dots\dots\dots (6)$$

This equation is known as the Stokes relationship. Applied to equation (4) it gives

$$-p_{ij} = -p_0 \delta_{ij} + \eta \left\{ \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3} \operatorname{div} u \right\}, \dots\dots\dots (7)$$

an expression in which only one coefficient of viscosity occurs. This expression has been taken over into most dynamical treatments of the subject of viscosity (see, for example, Glazebrook 3).

While the use of the Stokes relation (6) appears to be justified for monatomic gases, its use in other systems probably has no more justification than would the use of the Cauchy-Poisson relationship in all problems concerned with the elasticity of solids. Tisza(4), who has discussed the influence of  $\zeta$  on the absorption of supersonic waves (in which the medium is subjected to rapidly fluctuating bulk stresses), adduced evidence to show that, for polyatomic gases and liquids  $\kappa$  is a positive quantity greater than  $\eta$ . For such systems  $\zeta$  must also be a positive quantity.



Mandel'shtam and Leontovich(5, 6) have provided similar evidence from consideration of the absorption of ultrasonic waves in liquids. The implications of their findings have been discussed by Frenkel(7, pp. 208-11), who arrived at an equation

$$-\frac{1}{V} \frac{\partial V}{\partial t} = \frac{1}{K_2} \frac{d \Delta p_2}{dt} + \frac{1}{\eta_2} \Delta p_2, \dots\dots\dots (8)$$

with  $V$  the volume of the assembly,  $K_2$  the bulk modulus (not necessarily the same as the static value of  $K$ ),  $\eta_2$  the volume viscosity (corresponding to the  $\kappa$  of Stokes), and  $\Delta p_2$  the non-equilibrium part of the pressure change (the  $p - p_0$  of expression (5) above). Equation (5), which since it is concerned with steady pressures only, does not contain a term in  $dp/dt$  and in the bulk modulus, can be written

$$p - p_0 = \Delta p = \kappa \operatorname{div} u = -\frac{\kappa}{V} \frac{\partial V}{\partial t} = -\frac{\kappa}{\rho} \frac{\partial \rho}{\partial t}, \dots\dots\dots (9)$$

where  $\rho$  is the density of the fluid.

### III. MEASUREMENT OF VISCOSITY BY RATE OF EXTENSION

The measurement of the viscosity of a semi-solid mass (such as pitch) by measurement of the rate of extension of fibres, or the rate of compression of a cylinder under a given stress was introduced by Trouton(8), who studied the elongation of specimens of pitch and wax under tensile stresses. Trouton defined a coefficient of viscous traction ( $\lambda$ ) by

$$\lambda = \frac{\text{tensile stress}}{\text{tensile strain velocity}}, \dots\dots\dots (10)$$

and proceeded to resolve his tensile forces into shear and dilatation components. He argued that the continued elongation of the specimen could only arise from the action of the shear forces since the dilatation forces could only have an *initial* effect. Since the resolved shear stress is equal to one-third of the applied tensile stress it follows that:

$$\text{Ordinary viscosity } (\eta) = \frac{\text{tensile stress}}{3 \times \text{tensile strain velocity}},$$

$$\text{or} \quad \eta = \lambda/3 \dots\dots\dots (11)$$

This method has been used by many other workers(9-16). The results obtained have frequently been shown to be consistent with viscosities as calculated from torsional measurements by means of the expression:

$$\text{Viscosity} = \frac{\text{shear stress}}{\text{shear strain velocity}} \dots\dots\dots (12)$$

It has, however, not fully been realized that expressions (11) and (12) are strikingly inconsistent with the Stokes viscosity relationship.

Let us consider a system in which the  $p_{11}$  component alone is non-vanishing. In other words we are considering a pure tensile stress along the 1, or  $x$ , axis. Equation (4) now takes the form

$$p_0 - p_{11} = 2\eta \frac{\partial u_1}{\partial x_1} + \zeta \left( \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} \right),$$

$$0 = 2\eta \frac{\partial u_2}{\partial x_2} + \zeta \left( \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} \right),$$

$$0 = 2\eta \frac{\partial u_3}{\partial x_3} + \zeta \left( \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \frac{\partial u_3}{\partial x_3} \right).$$

On eliminating  $\frac{\partial u_2}{\partial x_2}$  and  $\frac{\partial u_3}{\partial x_3}$  from this set of equations we obtain

$$p_0 - p_{11} = \left( 2\eta + \frac{\eta\zeta}{\eta + \zeta} \right) \frac{\partial u_1}{\partial x_1} \dots \dots \dots (13)$$

For a system subject only to a pure torsion in the 1 direction,  $p_{23}$  is the only non-vanishing component of the stress tensor. Equation (4) then reduces to

$$-p_{23} = \eta \left( \frac{\partial u_2}{\partial x_3} + \frac{\partial u_3}{\partial x_2} \right) \dots \dots \dots (14)$$

Equation (14) is identical with equation (12). If in addition equation (13) is to be identical with equation (11), it follows that

$$2\eta + \frac{\eta\zeta}{\eta + \zeta} = 3\eta,$$

or

$$\zeta = \infty \dots \dots \dots (15)$$

This condition would also mean that  $\alpha = \infty$ . On the basis of equation (9) this means that  $\frac{\partial \rho}{\partial t} = 0$ , or the density does not undergo any progressive change under constant pressure. However, there are many rheological systems for which such a generalization is not necessarily true. Any material which can be tableted or briquetted, for example, progressively changes its density under a constant pressure. For such a system we therefore conclude that

$$\zeta \neq \infty,$$

and it must also follow that

$$p_0 - p_{11} \neq 3\eta \frac{\partial u_1}{\partial x_1},$$

so that for such systems the viscosity as calculated by means of equation (11) does not agree with that from equation (12). Several consistent differences between viscosity of tension and viscosity of torsion have been recorded in the literature. Thus Broome and Bilmes(17) have studied the extension of strips of bitumens and asphalts and have reported that the viscosity as calculated from such measurements is not identical with that from measurements in torsion.

Other investigators(18, 19, 20) have pointed out that a number of rheological systems show a progressive change in density when subject to tension. A specimen of asphalt studied by Lee, Reiner, and Ridgen(20) showed a total tensile extension of 7 per cent. and a density decrease of 2 per cent. Both the extension and density changes proceeded at an approximately constant rate from the moment of application of the load.

#### IV. BULK VISCOSITY AND ITS MEASUREMENT

The few examples quoted in the previous section indicate that while  $\zeta$  is usually a large quantity in rheological systems (and not a small negative quantity as postulated by the Stokes relationship) it is not always sufficient to take

$\zeta = \infty$ . Any substance for which  $\zeta \neq \infty$  will show a progressive change in density at a fixed hydraulic pressure. Such a progressive change in density is of common occurrence in briquetting procedure. In this operation a sample of the powder to be treated is placed in a mould and extreme pressure applied by means of a plunger. The primary effect of the plunger is to apply a single compressive stress component ( $-p_{11}$ ). Besides producing compression in the direction of the stroke, such a stress would cause the assembly to expand laterally. However, the walls of the mould prevent such lateral expansion and if we assume that the Poisson rate for the material is of the order, but not necessarily equal to, 0.5 it would follow that the walls would exert a normal stress component equal to  $-p_{11}$ . Houwink(21) has shown that most substances capable of sustaining large strains have values for the Poisson ratio very nearly 0.5. The material is therefore effectively subjected to an hydraulic pressure and the factor which determines the rate of briquetting is thus the quantity  $\kappa$  introduced above. This quantity may be referred to as the bulk viscosity since it measures the magnitude of the bulk stress required to produce a unit rate of deformation. Frenkel has called this same quantity the volume viscosity. It is related to the first and second coefficients of viscosity by equation (5).

The second coefficient  $\zeta$  is, however, not directly obtainable from rheological measurements and it is therefore more convenient to express  $\kappa$  in terms of the two readily measurable quantities:  $\eta$ , the coefficient of torsional viscosity (or ordinary viscosity) and  $\lambda$  the coefficient of viscous traction. We then get

$$\lambda = 2\eta + \frac{\eta\zeta}{\eta + \zeta}, \dots\dots\dots (16)$$

or

$$\zeta = \eta \frac{2\eta - \lambda}{\lambda - 3\eta} \dots\dots\dots (17)$$

It therefore follows that

$$3\kappa = \eta \frac{\lambda}{3\eta - \lambda},$$

or

$$\frac{1}{\lambda} = \frac{1}{3\eta} + \frac{1}{9\kappa}, \dots\dots\dots (18)$$

an expression for the relations between coefficients of viscosity exactly similar to the corresponding relation holding between moduli of elasticity (equation (1)). These expressions are similar simply because one is derived from consideration of a strain tensor and the other from a strain-velocity tensor, every term of which is a temporal derivative of a corresponding term in the former tensor. However, equation (18) has apparently not been given before and could with advantage be considered experimentally.

For an experimental test of equation (18) we require some system capable of being tableted or briquetted to yield a product which could subsequently be subjected to both torsional and tensile viscosity tests. Solid carbon dioxide was selected as suitable for triple tests of this nature. Carbon dioxide snow, after briquetting under hydraulic pressure, during which it undergoes a large density change, produces an extremely plastic solid.

Measurements of the bulk viscosity were obtained by pressing the snow into a mould by an hydraulic plunger under a fixed hydraulic pressure of



1300 lb./sq. in. or  $9.0 \times 10^7$  baryes. The times of pressing were varied and the density of each block formed was recorded. Figure 1 shows the densities obtained plotted against the times of pressing. The resultant curve is concave downwards, or the bulk viscosity increases as the density increases.

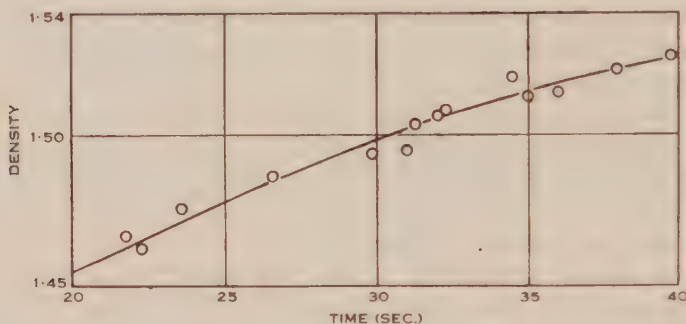


Fig. 1

At a density of  $\rho=1.50$  it may be seen that

$$\frac{1}{\rho} \frac{\partial \rho}{\partial t} = 0.0017,$$

from which it follows that, at  $\rho=1.50$ ,  $\alpha$  the coefficient of bulk viscosity is

$$\alpha = 5.4 \times 10^{10} \text{ poises.}$$

Samples were then cut from the 1.500 density blocks for tensile and torsion tests.

## V. FLEXURAL MEASUREMENT OF VISCOSITY

The bending of a beam was selected as the most suitable means of applying tensile forces to a sample of solid carbon dioxide. The samples produced as above were cut into slices roughly 20 cm. long by 6 cm. wide by 1 cm. thick. Each slice was placed over two parallel wooden supports of 1.2 cm. diameter and 15.6 cm. apart, and a wooden stirrup was placed centrally between them. Two small mirrors a measured distance apart were placed one at each end of the beam and light reflected from both was thrown onto a scale. The load on the stirrup was increased gradually until, at about 2 kg., the spot of light showed a progressive drift. The rate of movement, the total load including the weight of the stirrup, and the dimensions of the beam under test were recorded. The load was changed and another set of measurements obtained. This procedure was repeated until the beam fractured and it was then replaced by another cut from the same block. The stress in the outer fibres was computed from the recorded measurements on the usual assumptions made in deriving the modulus of rupture. The strain velocity in the outer fibres was also computed. Figure 2 shows the resultant strain velocity  $v$ . stress curve.

## VI. TORSIONAL MEASUREMENT OF VISCOSITY

Samples of the same batch of solid carbon dioxide were next cut in the form of blocks 20 by 6 by 6 cm. Each block as required was placed in a lathe and



the central 7 cm. turned down to a cylinder 3 cm. in diameter. The prepared block was then placed in a special torsion testing machine with fibre jaws of low thermal conductivity. One end was held fixed and a measured torque

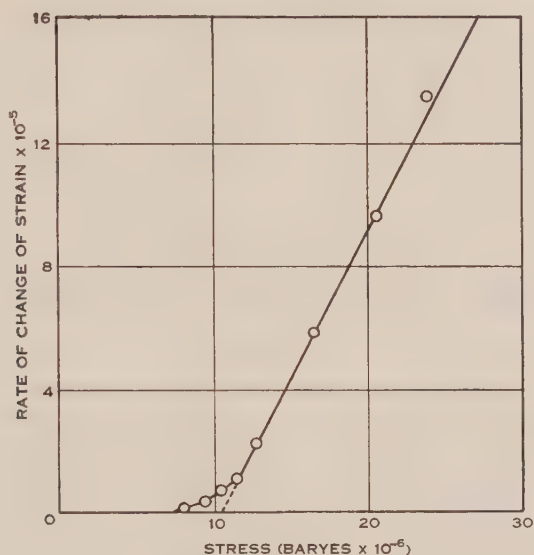


Fig. 2

applied to the other. Rotation of the free end was measured either with mirror and scale or, for larger angles, by protractor and pointer. Over a certain

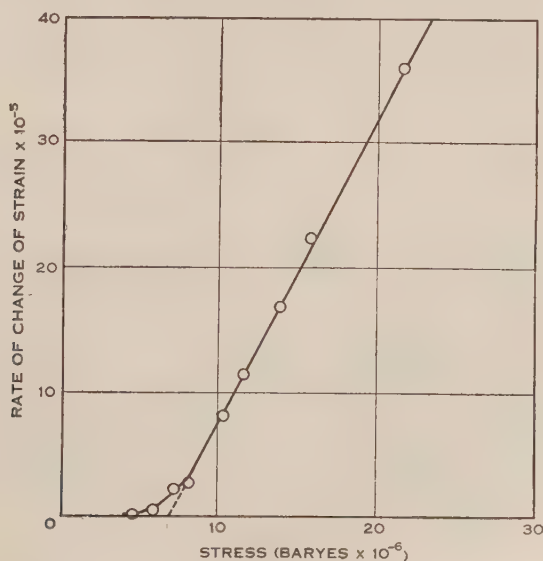


Fig. 3

range of torsional loading the free end turned at a constant rate and a total rotation of  $180^\circ$  or more was frequently recorded before the column fractured.

By taking measurements at a series of loads and by changing the block when fracture did occur, the rate of change of shear strain was obtained as a function of the shear stress. The results are plotted in Figure 3.

Figures 2 and 3 show essentially the same characteristics. At low stresses a fixed strain only is produced. At increasing stress there is a rapidly changing strain, and at higher stresses still, a large region for which there is a linear relationship between stress and rate of change of strain. Following the method of Bingham (Bingham and Green 22) we should define the viscosity as the slope of the linear portion of the stress *v.* strain velocity curve and the Bingham yield point (Houwink 20) as the intercept of the extrapolated linear portion of the graph. However, we require knowledge of the stresses in the outside fibres of a Bingham plastic exposed to flexural and torsional stresses. This information, while not directly available, may be derived quite easily.

#### VII. THE STRESS IN THE OUTSIDE FIBRES OF A BINGHAM PLASTIC EXPOSED TO FLEXURAL AND TORSIONAL STRESSES

Detailed calculation of the stresses in the outside fibres involves, as is well known, a knowledge of the variation of the stress throughout the cross-sectional area normal to the neutral axis, or line of zero stress. In the computation of the properties known as the modulus of rupture and torsional modulus of rupture respectively it is assumed that the body is Hookean in behaviour. This gives a value of  $2C/\pi R^3$  for the stress in the outside fibres of a body of radius  $R$  subjected to a torsional couple  $C$ . For a body subjected to a flexural bending moment  $M$  the corresponding figure for the stress in the outside fibres is  $6M/bd^2$ , where  $b$  is the breadth and  $d$  the depth of the rectangular specimen flexed. For substances which behave as Bingham plastics stresses less than some characteristic value  $f$  produce a zero strain velocity while stresses ( $\tau$ ) greater than  $f$  produce a strain velocity directly proportional to  $\tau - f$ . Such behaviour is very different from Hookean and it is clear that the curve shown in Figures 2 and 3 may not be directly interpreted as stress *v.* strain-velocity, or flow curves. The modifications required are, however, quite trivial, as may be seen by the arguments below. Consider first torsional loading.

(i) *Torsion Testing.*—Let us consider the application of an external couple  $C$  to one end of a cylindrical specimen of length  $L$ , and radius  $R$ . The other end is fixed. As a result let the free end turn through an angle  $\theta$  (which in the case of a plastic body will be a function of the time). At any fixed time the strain at the periphery is  $R\theta/L$ . For any fibre distance  $r$  from the axis of the couple the corresponding strain is  $r\theta/L$ . If  $D(-\theta/R/L)$  is the strain velocity in the outside fibres, that at a distance  $r$  from the axis of the couple is,  $D(r/R)$ . Now the (local) stress ( $\tau$ ) is related to the strain velocity by the Bingham equation

$$D \frac{r}{R} = \frac{1}{\eta} (\tau - f) H(\tau - f), \dots\dots\dots (19)$$

where  $H(x)$  is the Heaviside function and  $\eta$  is the viscosity. It appears at once that the axis of the couple, while a neutral strain-velocity axis, is only one of

minimum and not one of zero stress. Since  $\tau$  is never less than  $f$  we may ignore the Heaviside function in equation (19) and write for the external couple

$$C = \int_0^R 2\pi r^2 \tau dr,$$

$$\frac{C}{2\pi} = \int_0^R r^2 f dr + \int_0^R \frac{\eta D}{R} r^3 dr,$$

or 
$$\frac{C}{2\pi R^3} = \frac{1}{3}f + \frac{1}{4}\eta D, \dots\dots\dots (20)$$

or 
$$\frac{2C}{\pi R^3} = \eta D + \frac{4}{3}f \dots\dots\dots (20a)$$

Since Figure 3 plots  $2C/\pi R^3 v. D$  it follows that the slope represents the viscosity and the intercept  $4/3$  times the value of the Bingham yield point.

(ii) *Flexural Testing*.—In pure bending there is a central neutral plane (with respect to the strain) and a local strain (at any fixed time) proportional to the distance from the neutral plane. By following arguments similar to those used above we see that, for a loading which gives a steady strain-velocity, the local stress will attain a minimum value  $f$  on the neutral plane and elsewhere be proportional to the distance from that plane. Let  $y$  be the distance of any selected fibre from the neutral plane. Consider a rectangular beam of breadth  $b$  and depth  $d$ . The stress  $\tau$  for a Bingham system subject to tensile stress and having a coefficient of viscous traction  $\lambda$  is given by

$$\tau = f + 2\lambda y D/d, \dots\dots\dots (21)$$

where  $D$  is the tensile strain-velocity in the outside fibres. The total internal moment must balance the external moment  $M$ , so we have

$$M = 2 \int_0^{\frac{1}{2}d} bfy dy + 2 \int_0^{\frac{1}{2}d} \frac{2b\lambda D}{d} y^2 dy,$$

$$= \frac{bf d^2}{4} + \frac{b\lambda D d^2}{6},$$

or 
$$6M/bd^2 = \lambda D + 3f/2 \dots\dots\dots (22)$$

Since Figure 2 plots  $6M/bd^2 v. D$  it follows that the slope represents the coefficient of viscous traction and the intercept  $3/2$  times the value of the Bingham yield point.

## VIII. THE MAGNITUDE OF THE DIFFERENT VISCOSITY COEFFICIENTS FOR SOLID CARBON DIOXIDE

Application of equations (20a) and (22) to Figures 2 and 3 yields the following figures for the plastic properties of 1.500 density blocks of solid carbon dioxide :

### (i) *Under Shear Stress*

Viscosity ( $\eta$ )  $4.2 \times 10^{10}$  poises,

Bingham yield point  $4.8 \times 10^7$  baryes.

### (ii) *Under Tensile Stress*

Coefficient of viscous traction  $10.5 \times 10^{10}$  poises,

Bingham yield point  $6.9 \times 10^6$  poises.

An independent measurement of the bulk viscosity has given

Bulk (or volume) viscosity  $5.4 \times 10^{10}$  poises.

These three coefficients of viscosity, inasmuch as the theory of small strains can be applied to deformations of 5 per cent. or of that order, are connected by the expression given in equation (18). From the measured values of  $\eta$  and  $\lambda$  a calculated value of  $\kappa$  as  $7.0 \times 10^{10}$  poises may readily be obtained. This, in view of the limitations of both experimental and theoretical accuracy, is in satisfactory agreement with the  $5.4 \times 10^{10}$  poises from direct measurement and may be taken as evidence that the second viscosity coefficient, so long ignored in rheological discussions, has a definite experimental significance.

#### IX. REFERENCES

- (1) PLANCK, M.—“Mechanics of Deformable Bodies.” (Macmillan & Co.: London, 1932.)
- (2) STOKES, G. G.—“Mathematical and Physical Papers.” Vol. 1, pp. 82-8. (Cambridge Univ. Press: Cambridge, 1880.)
- (3) GLAZEBROOK, R.—“Dictionary of Applied Physics.” Vol. 1, p. 345. (Macmillan & Co.: London, 1922.)
- (4) TISZA, L.—*Phys. Rev.* **61**: 531-6 (1942).
- (5) MANDEL'SHTAM, L., and LEONTOVICH, M.—*C.R. Acad. Sci. U.R.S.S.* **3**: 111-4 (1936).
- (6) MANDEL'SHTAM, L., and LEONTOVICH, M.—*J. Exp. Theor. Phys.* (Russ.) **7**: 438 (1937).
- (7) FRENKEL, J.—“Kinetic Theory of Liquids.” pp. 208-13. (Oxford, 1946.)
- (8) TROUTON, F. T.—*Proc. Roy. Soc. A* **77**: 426-40 (1906).
- (9) TROUTON, F. T., and ANDREWS, E. S.—*Proc. Phys. Soc.* **19**: 47-57 (1909).
- (10) ANDRADE, E. N. DA C.—*Proc. Roy. Soc. A* **84**: 1-12 (1910).
- (11) ANDRADE, E. N. DA C.—*Phil. Mag.* **27**: 869-70 (1914).
- (12) ANDRADE, E. N. DA C., and CHAMBERS, B.—*Proc. Roy. Soc. A* **138**: 348-74 (1932).
- (13) ROLLER, P. S.—*J. Phys. Chem.* **43**: 457-89 (1939).
- (14) LILLIE, H. R.—*J. Amer. Ceram. Soc.* **16**: 619-31 (1933).
- (15) RICHMOND, W. O.—*Physics* **5**: 131 (1934).
- (16) STEINBERGER, R. L.—*Ibid.* **5**: 53-62 (1934).
- (17) BROOME, D. C., and BILMES, L.—*J. Soc. Chem. Ind.* **60**: 184-90 (1941).
- (18) GLANVILLE, W. H., and THOMAS, F. C.—Tech. Pap. Build. Res. Lond. No. 23. (H.M. S.O.: London, 1939.)
- (19) SWAINGER, K. H.—*Nature* **158**: 165 (1946).
- (20) LEE, A. R., REINER, M., and RIDGEN, P. J.—*Ibid.* **158**: 708 (1946).
- (21) HOUWINK, R.—“Elasticity, Plasticity, and the Structure of Matter.” (Cambridge Univ. Press: Cambridge, 1937.)
- (22) BINGHAM, E. C., and GREEN, H.—*Proc. Amer. Soc. Test. Mater.* **19**: 640-75 (1919).



# THE DIELECTRIC PROPERTIES OF BINARY SYSTEMS OF KETONES AND HYDROCARBONS

By R. J. MEAKINS\*

[Manuscript received April 13, 1949]

## Summary

Experimental work on the dielectric properties of solid systems of long-chain polar and non-polar compounds has been performed by Jackson(1), Sillars(2), and Pelmore(3) with dilute solutions of esters in paraffin wax. In subsequent theoretical studies in this field Fröhlich(4) adopted, as a model, a crystalline long-chain paraffin in which a small proportion of the molecules were assumed to be replaced by long-chain polar molecules. Taking Debye's theory of dielectric absorption in polar solids as a basis, this led to the derivation of a relationship between the relaxation time and the number of links in the polar molecule.

In order to obtain experimental results more closely related to the theory, electrical measurements have now been made with symmetrical long-chain ketones in pure crystalline hydrocarbon solvents. The variation of dielectric constant and loss angle with frequency for these systems is found to be approximately in agreement with the Debye theory of dielectric absorption.

For ketones of chain-length shorter than that of the hydrocarbon solvent the relationship between relaxation time and chain-length is of the type expected from Fröhlich's theory, but displaced towards lower frequencies. The results show that, in applying the theoretical relationship to systems of long-chain polar and non-polar compounds, it would be necessary to determine new constants for each different type of polar compound and for each different solvent.

The results of electrical measurements of ketone-hydrocarbon systems with ketones of varying chain-length (both less and greater than that of the solvent) are in accordance with conclusions from previous X-ray studies of long-chain hydrocarbons.

The effects of dipole interaction on the dielectric absorption of systems of laurone and *n*-hexacosane, are discussed in relation to the solid-liquid phase diagram. The results of dielectric measurements with pure crystalline compounds suggest that dipole interaction and electrical cooperative effects are larger in ketones than in esters of similar chain-length.

## I. INTRODUCTION

In determining the dipole moments of polar compounds, errors due to dipole interaction are avoided by making the appropriate electrical measurements with dilute solutions in non-polar solvents. This principle was adopted with solid solutions by Jackson(1) in an investigation of the electrical properties of paraffin wax to which a small proportion of a long-chain polar compound, cetyl palmitate, had been added. It was found that the variation of power loss with frequency and temperature for this solid solution were in good agreement with the Debye theory of dielectric absorption. The relaxation time,

\* Division of Industrial Chemistry, seconded to Division of Electrotechnology, C.S.I.R.O.

calculated from the frequency for the maximum loss angle indicated, according to Debye's theory, an apparent viscosity of the order of that of castor oil, and to explain this it was suggested that only the polar region of the cetyl palmitate molecule oscillates in an alternating electric field. Later experiments(2) with other long-chain esters showed, however, that the relaxation time depends on the number of links in the ester chain, indicating that orientation of the dipole involves movement of the molecule as a whole. This was confirmed(3) by the results of experiments with the dioctyl esters of azelaic and sebacic acids, in which the polar groups are oriented in the like and opposite senses respectively. The negligible dielectric absorption observed with the latter indicated that orientation of the dipoles does not occur independently of the main chain. In contrast, the former ester gave considerable dielectric loss.

In explaining the results described above, Fröhlich(4) suggested that, in accordance with Debye's theory of dielectric absorption in solids, the polar molecules in such solid solutions do not rotate freely but have two equilibrium positions of equal energy and opposite direction, with a certain probability of transition between them. The application of an electric field results in a polarized condition in which a slightly greater proportion of the dipoles occupy positions favouring the field. Thus, for ice, having a dielectric constant of 80, Debye(5) has calculated that a field intensity of 1 v./cm. would lead to the turning of 1 molecule in every five million. The time taken to approach equilibrium after the application or removal of the electric field is measured in terms of the relaxation time,  $\tau$ , which is defined by

$$\tau = \frac{1}{2p},$$

where  $p$  is the probability that an individual dipole will, in unit time, make the transition from one equilibrium position to another.

In a periodic field the polarization lags, leading to dielectric loss, the loss angle being given approximately by Debye's equation

$$\tan \delta = \frac{\epsilon_0 - \epsilon_\infty}{\sqrt{\epsilon_0 \epsilon_\infty}} \cdot \frac{\omega \tau}{1 + \omega^2 \tau^2}, \quad \dots \dots \dots (1)$$

where  $\epsilon_0$  is the dielectric constant for static fields,

$\epsilon_\infty$  is the dielectric constant for high frequency fields,

$\omega$  is the angular frequency of the alternating field.

This equation is the same as that derived for polar liquids in which the molecules are assumed to rotate in the electric field.  $\tan \delta$  reaches a maximum at an angular frequency given by  $\omega \tau = 1$  which provides a basis for the experimental determination of  $\tau$ .

The crystal structure of long-chain hydrocarbons has been extensively studied by Müller(6, 7) and by Bunn(8) using X-ray diffraction. The unit cell is usually rectangular, with the dimensions of the base approximately  $5.0 \text{ \AA} \times 7.5 \text{ \AA}$ , and the height slightly greater than the chain-length. The molecular chains are perpendicular to the rectangular base at the corners and at the intersection of the diagonals. In a theoretical study of the dielectric properties of systems of long-chain polar and non-polar compounds Fröhlich

adopted a model of this type, in which a small proportion of the hydrocarbon molecules were replaced by polar molecules. This led to the derivation of a relationship between the number of links,  $n$ , in the molecular chain of the polar compound and the relaxation time,  $\tau$ , assuming the chain-length of the polar compound to be less than that of the hydrocarbon

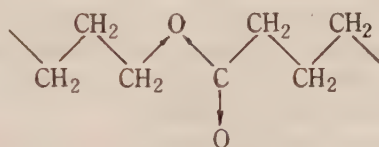
$$\log 2\pi\tau = \alpha + \frac{\pi^2 B^2 \bar{n}}{4kT} \cdot \tan h \frac{n}{2\bar{n}}, \dots\dots\dots (2)$$

where  $\alpha$ =a constant, determined from the experimental results of Sillars and Pelmore,

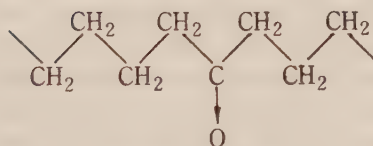
$B^2$ =the energy required per link to turn the carbon chain by an angle of  $\sqrt{2}$  radian,

$\bar{n}=A/B$ ,  $A^2$  being a constant related to the energy required to twist the chain, i.e. deform the valency angle,

$k$ =Boltzmann's constant.



ESTER LINKAGE



KETONE LINKAGE

Fig. 1.—Ester and ketone structures showing polar bonds.

The constants  $B^2$  and  $\bar{n}$  were estimated theoretically and also determined empirically from the experimental results, the values being in quite good agreement. Substitution of the numerical values for these and other constants in equation (2) gave the following:

$$\log 2\pi\tau = -50.4 + \frac{13800}{T} \cdot \tan h \frac{n}{26} \dots\dots\dots (3)$$

Since paraffin wax consists of a mixture of hydrocarbons of different molecular chain-length, and is therefore not a simple crystalline structure, Fröhlich's theory does not apply precisely to the experiments of Jackson, Sillars, and Pelmore. In view of this it seemed desirable to investigate the dielectric properties of solid solutions in pure crystalline hydrocarbons. For this programme ketones appeared to be more suitable polar solutes since they possess a single polar linkage at right angles to the main chain, whereas the ester dipole is more complex due to the presence of three polar linkages (see Fig. 1). Moreover, the crystal symmetry of the ketones is identical with that of the



hydrocarbons(9) whereas the ester molecules may be tilted with respect to the base of the unit cell(10). Symmetrical ketones were chosen chiefly because of ease of preparation.

## II. MATERIALS

All melting points are uncorrected.

The hydrocarbons *n*-docosane, *n*-hexacosane, and *n*-triacontane were prepared by the electrolysis of the potassium salts of the appropriate fatty acids(11). Those of intermediate chain-length, *n*-tetracosane and *n*-octacosane, were obtained as by-products in the preparation of the Grignard compounds from dodecyl and tetradecyl iodides respectively. With the hydrocarbons of chain-length up to 26 carbon atoms polar impurities were removed by recrystallization from alcohol, heating with concentrated sulphuric acid for 2 hours at 130 °C.(12), washing thoroughly with water, and recrystallization again from alcohol. The product was then tested by measuring its dielectric loss over the frequency range, 50 kc/s. to 50 Mc/s. and, if necessary, the purification process repeated. Hydrocarbons of chain-length greater than 26 carbon atoms were not sufficiently soluble in alcohol, and it was necessary to use light petroleum for their recrystallization.

The paraffin wax was a highly-refined commercial product, m.p. 63–65 °C., giving negligible dielectric loss in the frequency range of the Q-meter. This melting point indicates an average chain-length of about 29 carbon atoms.

The ketones were prepared from the appropriate fatty acids by the catalytic action of magnesium oxide at 330–360 °C.(13). The reaction product was treated with sulphuric acid to remove magnesium, and then with caustic soda to remove free fatty acid. This was followed by repeated washing with water, recrystallization from alcohol, and finally, vacuum distillation.

The long-chain esters used in this programme were prepared by heating together the silver salt of the appropriate fatty acid and the appropriate long-chain alkyl iodide.

The melting points of the hydrocarbons, ketones, and esters as found are given in Table 1.

For the measurement of the electrical properties of binary ketone-hydrocarbon mixtures, the required quantities of the ingredients were weighed (total c. 4–5 g.), melted together, and poured on a surface of mercury in a dish about 6 cm. in diameter, previously heated to a temperature a few degrees above the melting point of the wax mixture. The thick outer edge of the disk obtained on cooling was removed with a sharp microspatula and the cast then machined on a lathe fitted with a pneumatic chuck, to give flat surfaces.

## III. APPARATUS AND PROCEDURE

The measurements of loss angle and dielectric constant were made by a resonance method (Q-meter), using a sample-head similar to that described by Hartshorn and Ward(14). Readings at each frequency were taken first with the sample disk and then with the disk removed and the head adjusted to give the same capacitance. Small variations in the pressure of the capacitor plates



on the sample did not appreciably affect the final results. Measurements were made at each of nine different frequencies spaced at approximately equal logarithmic intervals within the range 60 kc/s.-50 Mc/s. For calculating the results at the latter frequency the residual parameters of the Q-meter measuring circuit were determined and the necessary corrections applied. It was possible, by the choice of suitable coils, to detect power factor increments of less than  $10^{-4}$  at frequencies up to about 10 Mc/s., the accuracy diminishing considerably above this point. Except where otherwise indicated, the electrical measurements were made in a laboratory conditioned at  $20 \pm 1^\circ \text{C}$ .

The melting points for the solid-liquid phase diagram were determined by the thaw-melt method(15).

TABLE 1  
MELTING POINTS OF HYDROCARBONS, KETONES, AND ESTERS

Hydrocarbons, Ketones, and Esters	Formulae	Melting Point ( $^\circ\text{C}$ .)
<i>n</i> -Docosane .. .. .	$\text{C}_{22}\text{H}_{46}$	44.4-44.8
<i>n</i> -Tetracosane .. .. .	$\text{C}_{24}\text{H}_{50}$	50.7-51.4
<i>n</i> -Hexacosane .. .. .	$\text{C}_{26}\text{H}_{54}$	56.4-56.7
<i>n</i> -Octacosane .. .. .	$\text{C}_{28}\text{H}_{58}$	61.2-61.6
<i>n</i> -Triacontane .. .. .	$\text{C}_{30}\text{H}_{62}$	65.7-66.5
8-Pentadecanone .. .. .	$\text{C}_{15}\text{H}_{30}\text{O}$	40.0-41.0
10-Nonadecanone .. .. .	$\text{C}_{19}\text{H}_{38}\text{O}$	58.4-58.9
11-Heneicosanone .. .. .	$\text{C}_{21}\text{H}_{42}\text{O}$	64.0-65.0
12-Tricosanone (laurone) .. .. .	$\text{C}_{23}\text{H}_{46}\text{O}$	68.3-68.8
13-Pentacosanone .. .. .	$\text{C}_{25}\text{H}_{50}\text{O}$	68.4-68.7
14-Heptacosanone (myristone) .. .. .	$\text{C}_{27}\text{H}_{54}\text{O}$	77.2-77.7
16-Hentriacontanone (palmitone) .. .. .	$\text{C}_{31}\text{H}_{62}\text{O}$	83.2-83.6
18-Pentatriacontanone (stearone) .. .. .	$\text{C}_{35}\text{H}_{70}\text{O}$	89.3-89.5
Dodecyl laurate .. .. .	$\text{C}_{24}\text{H}_{48}\text{O}_2$	29.3-29.8
Dodecyl myristate .. .. .	$\text{C}_{26}\text{H}_{52}\text{O}_2$	37.0-37.2
Cetyl palmitate .. .. .	$\text{C}_{32}\text{H}_{64}\text{O}_2$	52.4-53.2

#### IV. RESULTS AND DISCUSSION

##### (a) Dielectric Absorption and Dispersion of the Dielectric Constant for 0.1M Laurone in *n*-Hexacosane

It has been mentioned above that for a dilute solid solution of the long-chain ester cetyl palmitate in paraffin wax Jackson(1) found the variation of power factor ( $=\tan \delta$ , approx.) with frequency to be substantially in agreement with the Debye theory of dielectric absorption. However, the experimental curve was somewhat broader than the theoretical curve, calculated for the same maximum value of the power factor. Similar results were obtained in this Laboratory with ketone-hydrocarbon systems, this being illustrated in Figure 2 for 0.1M (approx. 3.4%) laurone in *n*-hexacosane. The broader nature of

the experimental curve is considered to be due to a spread of relaxation times about a principal value arising from slight differences in the environment of the polar molecules. The maximum loss angle for the system shown in Figure 2 occurs at a frequency of approximately 800 kc/s., giving a principal relaxation

$$\text{time } \tau = \frac{1}{2\pi f_{\max}} = 2.0 \times 10^{-7} \text{ seconds.}$$

i The anomalous dispersion of the dielectric constant, shown in Figure 2, is in approximate agreement with the Debye theory of dielectric absorption due to dipole relaxation. With the majority of the sample disks, however, the total variation was only slightly greater than the experimental error, and the dielectric constant was therefore not calculated from later measurements.

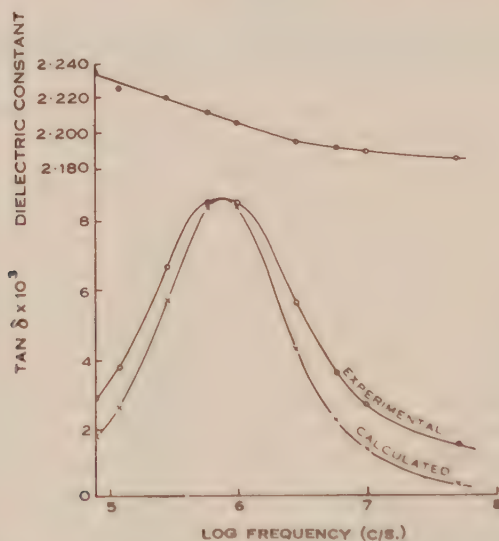


Fig. 2.—Variation of dielectric constant and power factor with frequency for 0.1M laurone in *n*-hexacosane at 20 °C.

(b) *Variation of the Relaxation Time due to Changes in Factors Affecting the Energy Barrier between Equilibrium Positions of the Polar Molecules*

In dilute solid solutions of long-chain polar compounds in paraffins, changes in the energy barrier between equilibrium positions of the polar molecules are most easily effected by varying the chain-length. The results of some experiments of this type are given in Figures 3, 4, and 5 showing the  $\tan \delta$ -frequency relationship for 0.1M solid solutions of a number of ketones in *n*-hexacosane, *n*-octocosane, and paraffin wax, m.p. 63–65 °C. In each case the frequency for the maximum  $\tan \delta$  decreases with increasing chain-length of the ketone, the relationship between  $\log f_{\max}$  and chain-length being approximately linear (see Fig. 6) for systems in which the ketone is of shorter chain-length than the solvent hydrocarbon. The theoretical curve, calculated from equation (3), shown in Figure 6, is also approximately linear but with a displacement towards

higher frequencies. Since equation (3) is in agreement with Sillars's results for esters in paraffin wax it is evident that the relaxation times of the ketones are longer than those of esters of similar chain-length and vary, moreover, according to the solvent. This indicates that in applying Fröhlich's relationship to systems of polar and non-polar compounds it is necessary to determine new constants for each different type of polar compound and for each different solvent(16).

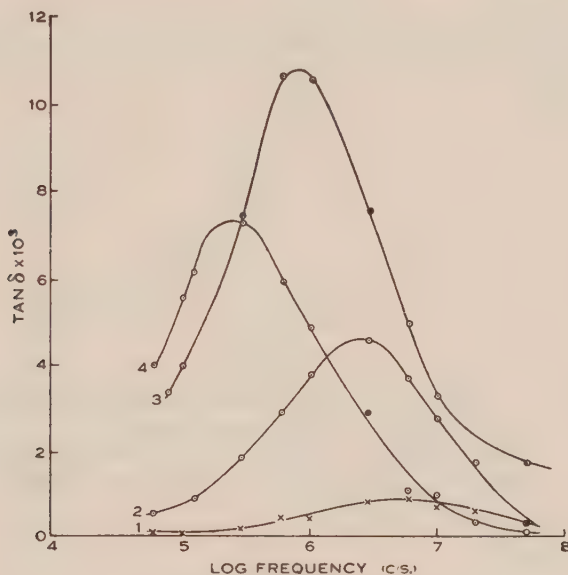


Fig. 3.—Dielectric absorption at 20 °C. for 0.1M solutions of ketones in *n*-hexacosane,  $C_{26}H_{54}$ .

- 1, 10-Nonadecanone,  $C_{19}H_{38}O$ ; 2, 11-Heneicosanone,  $C_{21}H_{42}O$ ;  
3, Laurone,  $C_{23}H_{46}O$ ; 4, 13-Pentacosanone,  $C_{25}H_{50}O$ .

With the *n*-hexacosane and *n*-octacosane systems, the relationship between  $\log f_{max}$  and chain-length shows a sharp discontinuity as the length of the ketone molecule exceeds that of the hydrocarbon. The reason for this is evident from the results of X-ray measurements by Piper *et al.*(12), who found that the spacing of the molecules in a crystalline hydrocarbon was very little affected by the addition of a small proportion of shorter molecules but was increased considerably by the presence of foreign molecules of larger chain-length. Thus, in the above-mentioned ketone-hydrocarbon systems the presence of ketone molecules longer than those of the hydrocarbon presumably expands the crystal lattice, leading to a smaller resistance to rotational transition, and hence giving shorter relaxation times. This effect is not obtained with paraffin wax, which is partly amorphous, and is therefore probably already an expanded lattice. In this solvent  $\log f_{max}$  tends towards a limiting value as the chain-length of the ketone increases, which is in agreement with Sillar's results for the corresponding ester systems. The expanded nature of paraffin wax presumably accounts for the shorter relaxation times observed for ketones in this solvent.

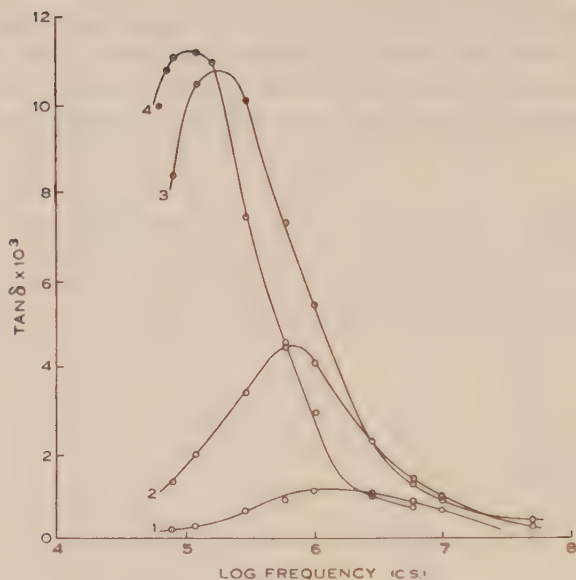


Fig. 4.—Dielectric absorption at 20 °C. for 0.1M solid solutions of ketones in *n*-octacosane,  $C_{28}H_{58}$ .  
 1, 11-Heneicosanone,  $C_{21}H_{42}O$ ; 2, Laurone,  $C_{13}H_{26}O$ ;  
 3, 13-Pentacosanone,  $C_{25}H_{50}O$ ; 4, Myristone,  $C_{17}H_{34}O$ .

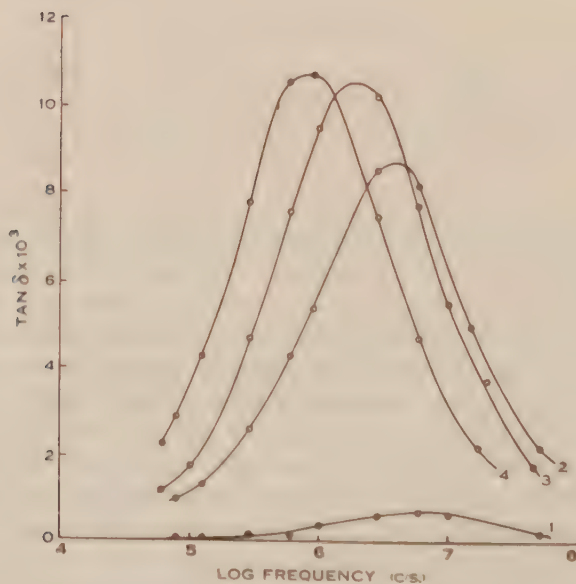


Fig. 5.—Dielectric absorption at 20 °C. for 0.1M solid solutions of ketones in paraffin wax, m.p. 63–65 °C.  
 1, 11-Heneicosanone,  $C_{21}H_{42}O$ ; 2, Laurone,  $C_{13}H_{26}O$ ;  
 3, 13-Pentacosanone,  $C_{25}H_{50}O$ ; 4, Myristone,  $C_{17}H_{34}O$ .



The results shown in Figure 6 indicate that the ketones give slightly longer relaxation times in *n*-octacosane than in *n*-hexacosane, suggesting that the resistance to rotational transition is slightly greater in the former solvent. This is in conformity with X-ray data(7) which shows that the closeness of packing of the molecules in crystalline hydrocarbons increases slightly with chain-length. The effect is further illustrated by the results given in Figures 7 and 8 for laurone and 10-nonadecanone systems. With laurone (Fig. 7) the maximum  $\tan \delta$  shifts slightly towards higher frequencies with each decrease in the chain-length of the solvent hydrocarbon between 30 and 24 carbon atoms. A further decrease to 22 carbon atoms brings the chain-length of the solvent

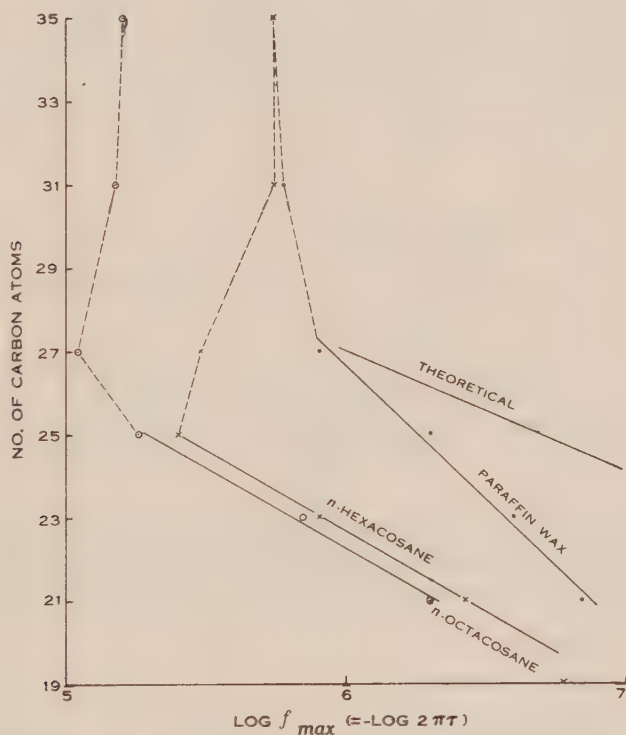


Fig. 6.—Variation of relaxation time with chain-length at 20 °C., for 0.1M solid solutions of symmetrical ketones in *n*-hexacosane, *n*-octacosane, and paraffin wax, m.p. 63–65 °C.

below that of the laurone, and here a large increase in  $f_{max}$  is observed. This is presumably due to the expansion of the hydrocarbon crystal lattice due to the presence of foreign molecules of longer chain-length (see above).

The two dielectric absorption curves given for 10-nonadecanone (Fig. 8) also indicate a higher  $f_{max}$  for the solvent of shorter chain-length.

Figure 7 shows that, with laurone systems, the dielectric absorption diminishes as the chain-length of the hydrocarbon solvent increases above 26 carbon atoms, the maximum  $\tan \delta$  values ( $\times 10^3$ ) being 10.9, 4.5, and 1.3

for solvents of chain-length 26, 28, and 30 carbon atoms respectively. Laurone systems with *n*-dotriacontane ( $C_{32}H_{66}$ ) and *n*-tetratriacontane ( $C_{34}H_{70}$ ) gave negligible dielectric loss in the frequency range of the Q-meter. With

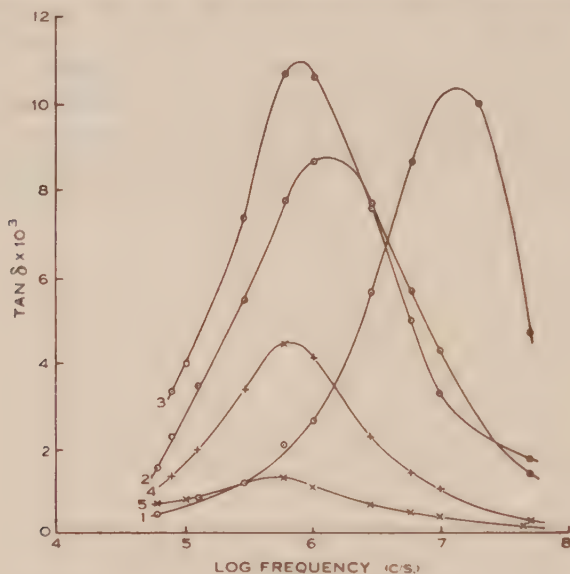


Fig. 7.—Dielectric absorption at 20 °C. for 0.1M solid solutions of laurone,  $C_{22}H_{44}O$ , in various hydrocarbons.

1, *n*-Docosane,  $C_{22}H_{46}$ ; 2, *n*-Tetracosane,  $C_{24}H_{50}$ ; 3, *n*-Hexacosane,  $C_{26}H_{54}$ ; 4, *n*-Octacosane,  $C_{28}H_{58}$ ; 5, *n*-Triacontane,  $C_{30}H_{62}$ .

10-nonadecanone (Fig. 8) also, it was found that the solvent of chain-length much longer than the ketone gave very little dielectric absorption compared with that obtained with the solvent more nearly equal in chain-length. Similar

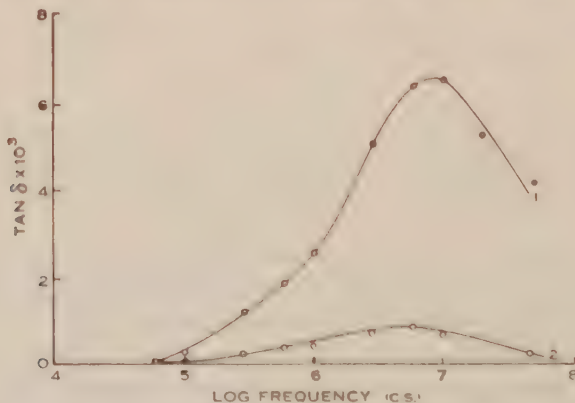


Fig. 8.—Dielectric absorption at 20 °C. for 0.1M solid solutions of 10-nonadecanone,  $C_{19}H_{38}O$ , in the hydrocarbons:

1, *n*-Docosane,  $C_{22}H_{46}$ ; 2, *n*-Hexacosane,  $C_{26}H_{54}$ .

results were obtained with 11-heneicosanone, the maximum  $\tan \delta$  for the *n*-hexacosane system (see Fig. 3) being  $4.6 \times 10^{-3}$ , and for the *n*-octacosane system (see Fig. 4)  $1.2 \times 10^{-3}$ . No explanation of this effect could be found at this stage, but a further investigation will be made when measuring facilities covering a wider frequency range become available.

(c) *Variation of Relaxation Time with Temperature for Dilute Ketone-Hydrocarbon Systems*

In polar solids, increasing thermal agitation causes an increased probability of transition of the molecules from one equilibrium position to another, leading to shorter relaxation times. The theoretical relationship, equation (2), shows that, for dilute solutions of long-chain polar compounds in paraffins,  $\log 2\pi\tau$  varies with temperature according to an inverse linear relationship. This is in agreement with the results of Sillars's experiments with ester systems, in which electrical measurements were made over the temperature range  $-45$  to  $+25$  °C.

TABLE 2  
RELATIONSHIP BETWEEN  $\log 2\pi\tau$  AND TEMPERATURE FOR DILUTE KETONE-HYDROCARBON SYSTEMS

System	Temperature (°C.)	$\log 2\pi\tau$
Laurone in <i>n</i> -hexacosane ..	20	$-5.9_0$
	30	$-6.2_1$
	40	$-6.8_1$
Myristone in <i>n</i> -octacosane ..	20	$-4.9_7$
	30	$-5.2_2$
	40	$-5.7_4$
Myristone in paraffin wax, m.p. 63-65 °C.	20	$-5.9_5$
	30	$-6.3_1$
	40	$-7.0_2$

In this Laboratory preliminary investigations with ketone systems were confined to the temperature range 20 to 40 °C., which could conveniently be obtained by operating the Q-meter in a heated box possessing suitable openings for the manipulation of the controls.

The values of  $\log 2\pi\tau$  at 20, 30, and 40 °C. for three ketone-hydrocarbon systems are given in Table 2.

From these results it is evident that, for the given temperature range, the relationship between  $\log 2\pi\tau$  and temperature is not linear, the effect of the rise from 30 to 40 °C. being much greater than that from 20 to 30 °C. This difference would be expected since, at the higher temperature in the present experiments, expansion of the lattice structure(7) takes place, causing a decrease

in the energy barrier between equilibrium positions, and a consequent additional decrease in  $\tau$ .

In a further investigation it is proposed to extend the range of temperature variation.

(d) *Effect of Dipole Interaction on the Dielectric Absorption of Ketone-Hydrocarbon Systems*

The investigations discussed in the preceding sections deal with dilute (0.1M) solid solutions, in which the majority of the ketone molecules are probably surrounded by non-polar solvent and dipole interaction is insignificant. Under these conditions, according to Fröhlich(4), the two positions of equilibrium of the polar molecules possess equal potential energy, rotational transitions from one position to the other leading to dielectric loss in an alternating electric field.

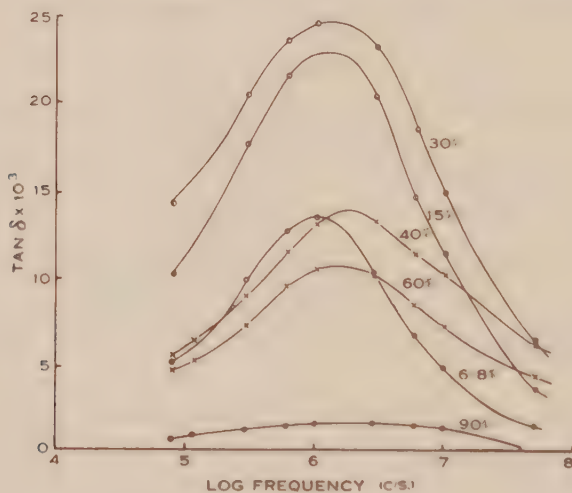


Fig. 9.—Dielectric absorption at 20 °C. for a number of different concentrations of laurone in *n*-hexacosane.

With concentrated solutions, or with pure crystalline compounds, in which dipole interaction is a significant factor, there may be a considerable energy difference  $V$  between equilibrium positions(17). At temperatures where  $kT \ll V$ , the probability of finding a turned dipole, and hence dielectric loss of the Debye type, will be very small.

An experimental investigation of the effect of dipole interaction on dielectric absorption was made(18) in this Laboratory for systems of laurone and *n*-hexacosane. The power factor-frequency relationships were determined for pure *n*-hexacosane and pure laurone, and for a number of different concentrations of laurone in *n*-hexacosane. The results, shown graphically in Figure 9, indicate that the dielectric absorption at first increases with increasing concentration, and then decreases, pure laurone giving negligible loss within the frequency range of measurement.



These results can be explained according to Fröhlich's theoretical models for dilute solid solutions and for pure crystalline polar compounds (4, 18, see above discussion). In the dilute solutions the majority of the laurone molecules are probably surrounded by non-polar solvent and therefore perform rotational transitions and contribute to the dielectric loss. A small proportion, however, occupies adjacent positions in the crystal lattice, and is subject to dipole interaction, which leads to a potential energy difference  $V$  between equilibrium positions. It is probable that for these molecules, even at room temperature  $kT \ll V$ , and their contribution to the dielectric loss is negligible. The proportion of adjacent molecules increases with increasing concentration, and the proportion contributing to the dielectric loss therefore decreases. This accounts for the rise in the dielectric absorption with increasing concentrations in the lower part of the range, and the fall to zero for pure laurone in which the polar molecules are all adjacent, and do not possess sufficient kinetic energy, at room temperature, to occupy the higher equilibrium position.

TABLE 3  
MAXIMUM TAN  $\delta$  AND FREQUENCY VALUES FOR CRYSTALLINE LONG-CHAIN  
ESTERS AT 20 °C.

Ester	Tan $\delta$ (max.)	Frequency (max.)
Methyl palmitate, $C_{17}H_{34}O_2$ .. ..	$1.0 \times 10^{-2}$	40 Mc/s.
Dodecyl laurate, $C_{24}H_{48}O_2$ .. ..	$1.4 \times 10^{-2}$	3 Mc/s.
Dodecyl myristate, $C_{26}H_{52}O_2$ .. ..	$4.0 \times 10^{-2}$	600 kc/s.
Cetyl palmitate, $C_{32}H_{64}O_2$ .. ..	$1.5 \times 10^{-2}$	100 kc/s.

With the more concentrated solutions of laurone in *n*-hexacosane, the absorption maxima occur at slightly higher frequencies than those of the more dilute solutions (see Fig. 9). This indicates a smaller energy barrier between equilibrium positions of the molecules, possibly owing to a partial solvent-solute reversal effect, the crystal lattice of the laurone being expanded as a result of the presence of hydrocarbon molecules of larger chain-length.

A more complete qualitative explanation of the mechanism of dielectric loss in systems of *n*-hexacosane and laurone is obtained from a study of the solid-liquid phase diagram (Fig. 10). This is of the "Eutectics of Solid Solutions" type, indicating that the two components are only partially soluble in the solid state. According to the concentration, systems of *n*-hexacosane and laurone may consist of either  $\alpha$ - or  $\beta$ -solid solutions, or of mixtures of these. The dielectric absorption is high for the  $\alpha$ -component and low for the  $\beta$ -component, as indicated above for dilute and concentrated solid solutions respectively. It therefore increases with concentration only in the region where the quantity

of laurone present as  $\alpha$ -solution continues to increase. At higher concentrations  $\beta$ -solution is formed at the expense of  $\alpha$ -solution and the dielectric absorption decreases.

With binary systems of this type the exact proportion of  $\alpha$ - and  $\beta$ -solid solutions present in the solid obtained from a particular melt depends greatly on the rate of cooling. If this is comparatively rapid, as is usual in the preparation of disks for dielectric measurements, equilibrium conditions are not attained immediately on solidification but are gradually established afterwards by molecular diffusion in the solid. It seemed probable therefore, that changes in the dielectric properties of such disks would occur as a result of prolonged

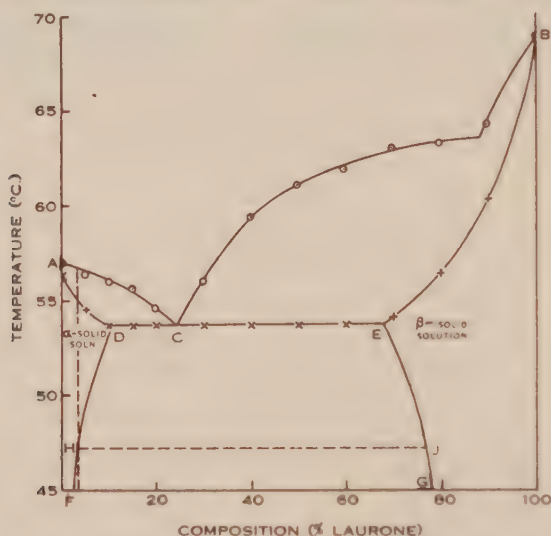


Fig. 10.—Solid-liquid phase diagram for the system *n*-hexacosane-laurone.

storage or on heating for shorter periods at elevated temperatures. This was found to be the case. The maximum  $\tan \delta$  value for a disk of 0.1M laurone in *n*-hexacosane was  $11.0 \times 10^{-3}$  immediately after preparation and  $9.6 \times 10^{-3}$  after storage for 20 days at 20°C. Heating for 82 hours at 45°C. caused a further reduction to  $7.3 \times 10^{-3}$ . On heating for 48 hours at 48.5°C., however, the  $\tan \delta$  (max.) increased to  $10.0 \times 10^{-3}$ . The power factor-frequency curves showing these changes are given in Figure 11.

The significance of these results is evident from the phase diagram. On cooling a 0.1M solid solution of laurone in *n*-hexacosane it becomes saturated with respect to laurone at a temperature represented by *H* on the solid-solid equilibrium curve *DF*, and, on further cooling molecular diffusion occurs and  $\beta$ -solid solution of composition represented by *J* is gradually formed. This process does not normally reach equilibrium, however, during the preparation of the disk, which initially consists almost entirely of  $\alpha$ -solid solution giving high dielectric loss. The gradual conversion of a proportion of the  $\alpha$ -solution to

$\beta$ -solution during storage, or on heating at 45 °C., leads to a reduction in the dielectric loss as indicated in Figure 11. After reaching equilibrium at 45 °C., a further rise in temperature results in the  $\alpha$ -solution crystals becoming unsaturated with respect to laurone. Molecules of the latter then diffuse into the  $\alpha$ -solution, this being accompanied by an increase in the dielectric absorption.

Changes in dielectric absorption were also observed on annealing the more concentrated *n*-hexacosane-laurone systems and also a 0.1M system of *n*-eicosan-10-one in *n*-hexacosane.

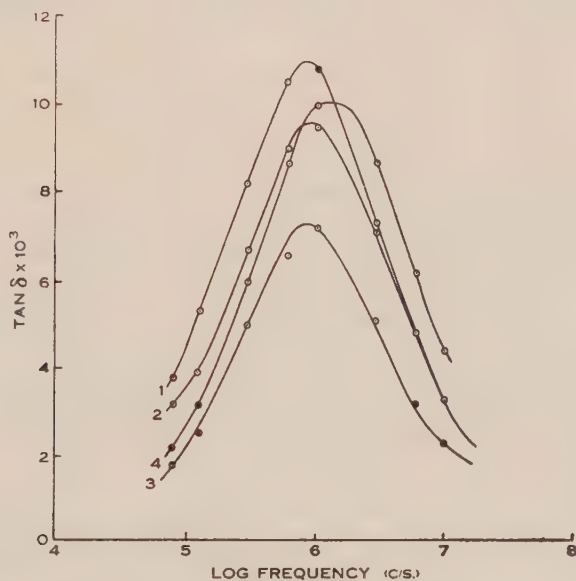


Fig. 11.—Effect of annealing on dielectric absorption of 0.1M laurone in *n*-hexacosane at 20 °C.

1, Freshly prepared disk ; 2, Disk stored 20 days at 20 °C. ; 3, Disk heated 82 hours at 45 °C. ; 4, Disk heated for a further 48 hours at 48.5 °C.

It has been stated above that several crystalline long-chain ketones gave negligible dielectric absorption, at 20 °C., in the frequency range 50 kc/s. to 50 Mc/s. In contrast, a number of long-chain esters gave strong absorption curves of the Debye type(18). The maximum  $\tan \delta$  and frequency values for three substances of this type, together with the values recently obtained by Dryden and Jackson(19) for the ester methyl palmitate, are given in Table 3. It is evident that in these compounds rotational transitions of the molecules from one position of equilibrium to another take place at ordinary temperatures. This suggests that dipole interaction is less for the esters than for the ketones of corresponding chain-length.

Although crystalline ketones gave negligible dielectric absorption in the frequency range of the Q-meter, significant loss was observed at 1000 c/s., the loss angle increasing with decreasing chain-length. This is indicated by the

results given in Table 4. In further experiments it was found that the loss angle for a particular ketone varied considerably (by a factor of at least 3) according to the rate of cooling during crystallization and the period of storage after preparation.

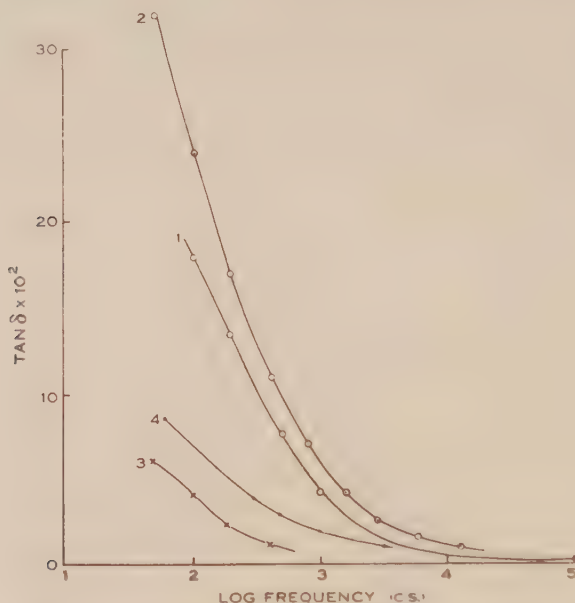


Fig. 12.—Low frequency dielectric absorption at 20 °C. for.  
1, 8-Pentadecanone,  $C_{15}H_{30}O$ ; 2, 9-Heptadecanone,  $C_{17}H_{34}O$ ;  
3, Cetyl palmitate,  $C_{32}H_{64}O_2$ ; 4, Methyl palmitate,  $C_{17}H_{34}O_2$ (19).

It has been suggested above that in pure crystalline ketones the rotational transitions of the individual dipoles from one position of equilibrium to another are suppressed as a result of dipole interaction. It is possible, however, that groups of dipoles are able to reverse, commencing at crystal faces or lattice defects. Such electrical cooperative effects(20) would be expected to lead to

TABLE 4  
TAN  $\delta$  AT 1000 C/S. FOR SYMMETRICAL LONG-CHAIN KETONES

Ketone	Tan $\delta$
8-Pentadecanone, $C_{15}H_{30}O$ .. .. .	$6 \times 10^{-3}$
10-Nonadecanone, $C_{19}H_{38}O$ .. .. .	$3 \times 10^{-3}$
12-Tricosanone (laurone), $C_{23}H_{46}O$ .. .. .	$5 \times 10^{-3}$
14-Heptacosanone (myristone), $C_{27}H_{54}O$ .. .. .	$2 \times 10^{-3}$
16-Hentriacontanone (palmitone), $C_{31}H_{62}O$ .. .. .	$9 \times 10^{-4}$
18-Pentatriacontanone (stearone), $C_{35}H_{70}O$ .. .. .	$3 \times 10^{-4}$



high dielectric absorption at low frequencies, as observed for the crystalline ketones shown in Table 4. The smaller loss angle values obtained with slowly cooled, or stored, samples are probably due to the more perfect nature of the crystal lattice.

The results of some further measurements of the low frequency dielectric absorption for two ketones and an ester, together with those obtained by Dryden and Jackson for methyl palmitate(19), are shown graphically in Figure 12.\* In the esters, dielectric absorption regions occur both at radio and audio frequencies, suggesting that both individual and group reversal of the dipoles takes place.

This programme is being continued.

#### V. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Electrotechnology, C.S.I.R.O. The author wishes to thank Professor D. M. Myers,† University of Sydney, for his interest in, and encouragement of, this programme, and also Mr. A. M. Thompson‡ for his cooperation in setting up the electrical apparatus required. Acknowledgment is also due to Mr. A. M. Thompson,‡ Mr. R. K. Oliver,¶ and Mr. J. M. Melano¶ for taking the measurements at low frequencies, and to Mr. J. S. O'Rourke¶ for help in the use of the Q-meter. Comments on this paper by Professor D. M. Myers, Dr. R. A. Sack,§ and Dr. L. G. Rees§ were greatly appreciated.

#### VI. REFERENCES

- (1) JACKSON, W.—The mechanism of dielectric loss in paraffin wax solutions at high radio frequencies. *Proc. Roy. Soc. A* **150**: 197 (1935).
- (2) SILLARS, R. W.—The behaviour of polar molecules in solid paraffin wax. *Ibid.* **169**: 66 (1939).
- (3) PELMORE, D. R.—Dielectric loss due to polar molecules in solid paraffin wax. *Ibid.* **172**: 502 (1939).
- (4) FRÖHLICH, H.—Dielectric loss in paraffin wax solutions. *Proc. Phys. Soc.* **54**: 422 (1942).
- (5) DEBYE, P.—“Polar Molecules”. p. 106. (The Chemical Catalog Co. Inc.: New York, 1929.)
- (6) MÜLLER, A.—The crystal structure of normal paraffins at temperatures ranging from that of liquid air to their melting points. *Proc. Roy. Soc. A.* **127**: 417 (1930).
- (7) MÜLLER, A.—An X-ray investigation of normal paraffins near their melting points. *Ibid.* **138**: 514 (1932).

\* Recent measurements in this Laboratory have shown that the loss at audio frequencies may be reduced considerably by recrystallization of the vacuum distilled product. This suggests that the loss may be due to small amounts of impurities and is not an intrinsic property of the compounds themselves.

† Formerly Chief, Division of Electrotechnology, C.S.I.R.O.

‡ Division of Electrotechnology, C.S.I.R.O.

¶ Formerly Division of Electrotechnology, C.S.I.R.O.

§ Division of Industrial Chemistry, C.S.I.R.O.

- (8) BUNN, C. W.—The crystal structure of long-chain normal paraffin hydrocarbons. The shape of the  $\text{CH}_2$  group. *Trans. Faraday Soc.* **35**: 482 (1939).
- (9) MÜLLER, A.—The dielectric polarization of *n*-long chain ketones near their melting points. *Proc. Roy. Soc. A* **158**: 403 (1937).
- (10) MALKIN, T.—Alternation in long-chain compounds. New X-ray data for long-chain ethyl and methyl esters and iodides, and a preliminary thermal examination of the esters. *J. Chem. Soc.* **1931**: 2796 (1931).
- (11) PETERSEN, J.—Electrolysis of the alkali salts of organic acids. *Z. Elektrochem.* **12**: 141 (1906).
- (12) PIPER, S. H., CHIBNALL, A. C., HOPKINS, S. J., POLLARD, A., SMITH, J. A. B., and WILLIAMS, E. F.—Synthesis and crystal spacings of certain long-chain paraffins, ketones and secondary alcohols. *Biochem. J.* **25**: 2072 (1931).
- (13) CURTIS, R. G., DOBSON, A. G., and HATT, H. H.—The ketonization of higher fatty acids with some observations on the mechanism of the reaction. *J. Soc. Chem. Ind.* **66**: 402 (1947).
- (14) HARTSHORN, L., and WARD, W. H.—The measurement of the permittivity and power factor of dielectrics at frequencies from  $10^4$  to  $10^8$  cycles per second. *J. Instn. Elect. Engrs.* **79**: 597 (1936).
- (15) GLASSTONE, S.—“Text-Book of Physical Chemistry”. p. 739. (Van Nostrand Co. Inc.: New York, 1940.)
- (16) MEAKINS, R. J.—The dielectric relaxation times of solid solutions of aliphatic long-chain esters and ketones in paraffin wax and *n*-hexacosane. *Nature* **162**: 994 (1948).
- (17) FRÖHLICH, H.—Dipolar interaction. *Trans. Faraday Soc.* **A 42**: 3 (1946).
- (18) MEAKINS, R. J.—Dielectric absorption in crystalline long-chain ketones. *Nature* **163**: 840 (1949).
- (19) DRYDEN, J. S., and JACKSON, W.—Dielectric behaviour of methyl palmitate. Evidence of resonance absorption. *Nature* **162**: 656 (1948).
- (20) FRANK, F. C.—Some mathematical models representing polar molecules in crystals. *Trans. Faraday Soc.* **A 42**: 24 (1946).

# ALKALOIDS OF THE AUSTRALIAN RUTACEAE: *ACRONYCHIA BAUERI*

## I. THE ISOLATION OF THE ALKALOIDS

By F. N. LAHEY\* and W. C. THOMAS

[Manuscript received March 19, 1949]

### Summary

The bark of *Acronychia Baueri* Schott yields four alkaloids, three of which, melicopine, melicopidine, and acronycidine, occur in *Melicope fareana*(1). The fourth is a new alkaloid *acronycine*,  $C_{20}H_{19}O_3N$ . The leaves yield a fifth alkaloid, melicopicine, also present in *Melicope fareana*.

## I. INTRODUCTION

An outline of a proposed extensive programme of alkaloidal research in Australia has been given by Price(1). Preliminary work by Price showed that one species of *Acronychia*, *A. Baueri* Schott, was rich in alkaloids and at his suggestion one of us (F.N.L.) has undertaken the investigation described in this series of papers.

*A. Baueri* is found in eastern Australia from Illawarra in New South Wales to Atherton in north Queensland. It has also been recorded from Lord Howe Island, the New Hebrides, and New Caledonia. It is a small tree which is confined to rain-forests. The inner bark, when cut, is a vivid yellow. An herbarium specimen, collected in 1824 by Allan Cunningham at Illawarra, New South Wales, gave good precipitates with alkaloidal reagents when tested recently by Webb (personal communication).

The species was placed by Engler and Prantl(2) under *Baurella australiana* Borzi, and belongs to the family Rutaceae, subfamily Toddaliodeae, subtribe Toddaliinae. This subtribe also includes the genera *Casimiroa*, *Toddalia*, and *Skimmia*, all known to yield alkaloids. The examination of the alkaloids of this species is now practically completed. In this, the first paper of the series, an account of the isolation and characterization of the alkaloids in the bark and leaves is presented.

A crude mixture of alkaloids is readily obtained by extraction of the bark with methanol followed by concentration of the solution. Four alkaloids have been shown to be present, of which melicopine, melicopidine, and acronycidine have been isolated simultaneously by Price (loc. cit.) from *Melicope fareana* but the fourth, for which we propose the name *acronycine*, does not appear to have been found previously. The isolation of each of these in a pure

\* Present address : Chemistry Department, University of Queensland.



state is a relatively easy matter. Melicopine is the least soluble in the usual organic solvents, particularly ether, and is separated readily from the others by crystallization. Acronycidine is the only one sufficiently basic to be extracted from a chloroform solution by 5 per cent. hydrochloric acid. Melicopidine and acronycine are precipitated as hydrochlorides from acetone; the former partially and the latter almost quantitatively. These are readily separated by crystallization from 10 per cent. hydrochloric acid from which only acronycine hydrochloride separates. A method of separation based on these observations is set out in the experimental section. The total yield of alkaloids ranges from 2.5-3 per cent. of the weight of dry bark. Melicopine is present to the extent of 1.3 per cent., acronycine 1-1.2 per cent., melicopidine 0.2-0.3 per cent., and acronycidine 0.1-0.2 per cent.

Acronycine separates from alcohol in shining yellow needles which sinter on drying and melt at 175-176 °C. A dilute solution of the base in alcohol is yellow with a bright green fluorescence. It forms an orange picrate, m.p. 150-154 °C., a red hydrochloride which separates from 10 per cent. hydrochloric acid in beautiful red needles, m.p. 125-130 °C. (decomp.), and a sulphate which crystallizes in long red needles from alcoholic sulphuric acid and melts at 158-159 °C. Analytical data indicate a molecular formula  $C_{20}H_{19}O_4N$  and the presence of one methoxyl group. Reagents for hydroxyl and carbonyl groups fail to give any reaction with the alkaloid.

None of the alkaloids present in the bark was found in the leaves but a fifth alkaloid, melicopine, isolated by Price (loc. cit.) from the bark of *Melicope fareana* was found in the leaves of *A. Baueri*. The leaves also contain a mixture of more strongly basic alkaloids but this has not yet been examined.

## II. EXPERIMENTAL

All melting points are corrected unless otherwise stated. Microanalyses are by Dr. K. Tettweiler and Mr. N. L. Lottkowitz.

### (a) Extraction of the Alkaloids

The milled dry bark (15 kg.) was extracted thoroughly in a Soxhlet with methanol. The extract (approximately 10 l.) after standing overnight was filtered, the residue being mainly crude triterpenoid material which will be the subject of a separate paper. Further quantities of this triterpene were obtained by concentrating the methanol extract to about half volume and again allowing to stand for 12 hours. The resultant precipitate contained triterpene mixed with some yellow alkaloidal material, which was removed by triturating the mixture with cold 10% hydrochloric acid and filtering. The treatment of this acid solution of the alkaloid will be dealt with later.

The filtered methanol extract was then concentrated to a syrupy consistency when it was dissolved in chloroform (3 l.), transferred to a separating funnel, and washed three times with 10% caustic soda solution (500 ml.). This extract on acidification gave a black tannin-like precipitate which was not further examined and the aqueous filtrate likewise yielded no tractable material. The chloroform solution after washing with alkali was extracted repeatedly with 5% hydrochloric acid. The combined hydrochloric acid extracts were made ammoniacal and extracted repeatedly with chloroform. The chloroform solution was dried with sodium sulphate, filtered, and the solvent completely removed leaving a dark crystalline mass. After repeated crystallization from methanol it separated in white needles or plates, m.p. 136.5°-137.5 °C.

The name acronycidine was given to this alkaloid. It was identical (m.p. and mixed m.p.) with the alkaloid isolated in the same manner simultaneously by Price(1) from *Melicope fareana*.



Found: C, 62.0; H, 5.6; N, 4.8%.

Calculated for  $C_{15}H_{15}O_5N$ : C, 62.3; H, 5.2; N, 4.8%.

The yield of acronycidine was found to vary greatly with different samples of bark. Yields ranging from 0.1 to 0.2% of the weight of dry bark were obtained.

The mother liquors from the recrystallizations of acronycidine contained a small amount of a steam-volatile base with an unpleasant smell. This rapidly deteriorated in the air and was not examined.

The original chloroform solution of the total alkaloids after extraction with 5% hydrochloric acid was dried over anhydrous sodium carbonate, filtered, and concentrated to a thick syrup which still contained some chloroform. To this were added approximately five volumes of ether. A quantity of tarry matter precipitated out quickly and the ether solution was decanted from this. More ether was added and, on standing, a crop of orange-yellow crystals separated which were filtered off and weighed (150 g.). If the hot extraction of the bark was replaced by cold percolation with methanol, the amount of tarry material was reduced, and the alkaloids were more easily obtained in a pure condition. However, quantitative removal of the alkaloids from the bark using the cold extraction method was extremely slow.

A further small crop (8 g.) of the same yellow alkaloid was obtained by distilling off all the solvent from the filtrate and repeating the precipitation with ether. The yellow alkaloid precipitated with the triterpene from the original methanol solution and separated by solution of the alkaloid in hydrochloric acid was recovered by dilution and treatment with ammonia. This material (35 g.) was found to be identical with the alkaloid precipitated by ether. The combined samples were recrystallized from chloroform-alcohol giving bright yellow needles, m.p. 178.5–179.5 °C. This alkaloid was identical (mixed m.p.) with melicopine from *Melicope fareana*(1). The yield of melicopine was approximately 1.3% of the weight of dry bark.

#### (b) The Ether-Soluble Alkaloids

The ether solution after removal of the melicopine was concentrated to a thick syrup and the residue dissolved in five or six times its volume of acetone. A small aliquot (100 ml.) of this solution was taken and concentrated hydrochloric acid added dropwise until no further precipitation of a red hydrochloride occurred. The remainder of the acetone was then treated with the corresponding amount of acid. This procedure saved considerable time as filtration of the very bulky precipitate was always necessary before the end point could be determined. The bright red hydrochloride was filtered and drained well at the pump. It was then dissolved quickly in boiling 10% hydrochloric acid (3 l.), the solution filtered from some tarry material with the aid of a little charcoal, and cooled rapidly in ice when a mass of red crystals separated. These were filtered, washed with cold 10% hydrochloric acid, and converted to the base by treatment with dilute ammonium hydroxide. The yellow alkaloid acronycine was filtered and crystallized several times from alcohol from which it separated in shining yellow needles which lost their lustre on drying and melted at 175–176 °C. The yield of acronycine varied from 1–1.2%.

Found: C, 74.9, 74.8; H, 6.0, 6.1; N, 4.3, 4.4\*;  $CH_3O$ , 9.6;  $CH_3N$ , 8.7%; mol. wt. 320 (Signer).

Calculated for  $C_{20}H_{19}O_3N$ : C, 74.8; H, 5.9; N, 4.4;  $CH_3O$ , 9.7 (one methoxyl);  $CH_3N$ , 9.0% (one methylimino); mol. wt. 321.

Acronycine is a weak base readily soluble in chloroform, ether, benzene, and acetone and is best crystallized from alcohol or ethyl acetate. A dilute solution in alcohol is pale yellow with a green fluorescence. Its hydrochloride forms bright red needles from 10% hydrochloric acid. However, this salt could not be dried satisfactorily for analysis as hydrogen chloride is gradually given off even at room temperature. Its melting or decomposition point depends largely on

\* Nitrogen estimations on acronycine and its derivatives were unsatisfactory by the usual Dumas method. We are greatly indebted to Dr. K. Tettweiler for introducing to these laboratories a new procedure due to Unterzaucher (unpublished) in which the combustion is done in the presence of oxygen and a specially active copper and which gave very satisfactory results.

the rate of heating but is usually in the range 125–130 °C. Acronyine forms an orange *picrate*, m.p. 150–154 °C.

Found: C, 56.2; H, 4.1; N, 10.8%.

Calculated for  $C_{36}H_{22}O_{19}N_4$ : C, 56.7; H, 4.0; N, 10.2%.

*Acronyine sulphate* separates in beautiful long red needles from alcohol containing concentrated sulphuric acid. It melts at 158–159 °C. (decomp.).

Acronyine on treatment in turn with semicarbazide, 2,4-dinitrophenylhydrazine, acetic anhydride, and benzoyl chloride failed to give any derivative. The standard test for a methylenedioxy group likewise was negative.

Two mother liquors remained from the isolation of acronyine, one the acetone solution after precipitation of acronyine hydrochloride and the other the 10% hydrochloric acid filtrates from recrystallization of the hydrochloride. Both of these solutions were diluted largely with water and extracted with chloroform. Both chloroform extracts were combined, dried with sodium sulphate, filtered, and the solvent removed. The residue was dissolved in alcohol and allowed to crystallize. A crop of yellow crystals separated which were filtered and dried. This alkaloid after several crystallizations melted at 121–122 °C. and was identical (mixed m.p.) with melicopidine from *Melicope fareana*(1). Further crops of melicopidine were obtained by concentrating the mother liquor. When the solution no longer gave reasonably pure melicopidine the solvent was completely removed and the residue treated with cold 10% hydrochloric acid and filtered with the aid of charcoal. The filtrate after dilution and basifying with ammonia, gave more melicopidine which was readily obtained pure by further crystallization from alcohol. The yield of melicopidine varied from 0.2–0.3% and the total yield of alkaloids lay between 2.6 and 3% of the weight of dry bark. No trace of the fourth alkaloid, melicopine, isolated by Price (loc. cit.) was found in this bark.

#### (c) *Extraction of the Leaves of Acronychia Baueri*

The dried milled leaves (1.8 kg.) were extracted with methanol by cold percolation. The solvent was completely removed from the extract and the residue treated with warm 5% hydrochloric acid (1 l.) followed by 10% hydrochloric acid (500 ml.). The acid solutions were combined after separation from the green sludge and extracted thoroughly with chloroform. This chloroform extract was dried with sodium sulphate, filtered, and distilled. The residue was again treated with cold 10% hydrochloric acid and the acid solution separated from a small amount of chlorophyll. This acid solution was diluted with an equal volume of water and again extracted with chloroform. After distilling off the chloroform the residue was dissolved in cold 10% hydrochloric acid and diluted till the acid strength was approximately 4%. On standing overnight bright yellow crystals separated. These were filtered and recrystallized from alcohol from which they separated in squat prisms, m.p. 133–134 °C., identical (mixed m.p.) with melicopine from *Melicope fareana*(1). The yield was between 0.3 and 0.4%. No other alkaloid could be detected in this fraction.

The original acid solutions after extractions with chloroform were basified with ammonia and again extracted thoroughly with chloroform. The bases so extracted are at present under investigation and will be the subject of a later paper.

### III. ACKNOWLEDGMENTS

The work described in this paper was carried out in the Chemistry Department, University of Melbourne. We are greatly indebted to Mr. L. J. Webb, C.S.L.R.O., for a supply of plant material, and to Mr. W. Nicklin, Mt. Glorious, Queensland, for permission to collect material from his property.

### IV. REFERENCES

- (1) PRICE, J. R.—*Aust. J. Sci. Res. A* 2: 249 (1949).
- (2) ENGLER, A., and PRANTL, K. "Die natürlichen Pflanzenfamilien." Vol. 19a. (Wilhelm Engelmann: Leipzig, 1931.)

# THE ISOLATION OF *d*-SPARTEINE FROM *HOVEA* SPECIES\*

By J. H. MORRISON† and K. G. NEILL‡

[Manuscript received March 28, 1949]

## Summary

The fairly widespread alkaloid *d*-sparteine has been isolated from two *Hovea* species. The constants of four salts prepared from it agreed with those previously recorded.

## I. INTRODUCTION

The genus *Hovea* which belongs to the family Leguminosae, tribe Genisteae, is entirely Australian, some eleven species being found, six of which occur in Queensland. As far as the authors are aware no previous chemical examination of this genus has been made. *H. longifolia* R.Br. occurs in several distinct forms, one of which shades into *H. acutifolia* A. Cunn., and it has been suspected of poisoning stock(1). Both of the above species gave positive alkaloid tests (Webb, personal communication) and were found to contain the steam volatile liquid alkaloid *d*-sparteine which has been previously isolated from several sources (cf. 2, 3, 4).

## II. EXPERIMENTAL

### (a) *Hovea acutifolia* A. Cunn.

This sample was collected at Walton Bridge near Brisbane, Queensland, in July 1947 and extraction begun shortly after drying.

The dried milled leaves were extracted with methanol in a soxhlet, and after removal of the solvent, the residue was treated with 1% hydrochloric acid, and filtered. The filtrate was made alkaline with caustic soda, steam distilled, and the liquid alkaloid recovered by ether extraction and distillation *in vacuo*.

It was obtained as a colourless, viscous liquid, yield 0.5%, b.p. 134 °C./1.1 mm. (lit.(5), 133–135 °C./1 mm.);  $[\alpha]_D^{20}$ , 16.8°; *c*, 7.2 in alcohol (lit.(6), 16.3°, *c*, 2 in alcohol).

Found: C, 76.4; H, 11.1; N, 11.8%.

Calculated for  $C_{15}H_{26}N_2$ : C, 76.9; H, 11.1; N, 12.0%.

### (i) *d*-Sparteine Dipicrate

Prepared in alcohol and recrystallized from a large volume of the same solvent as needles, m.p. 203–204.5 °C. (lit.(3), m.p. 204–205 °C.).

Found: N, 15.9%.

Calculated for  $C_{27}H_{32}N_8O_{14}$ : N, 16.2%.

### (ii) *d*-Sparteine Methiodide

Prepared in benzene and recrystallized from chloroform-benzene as colourless prisms, m.p. 234–235 °C. (lit.(6), m.p. 236–238 °C.).

\* The plant material was collected by L. J. Webb, C.S.I.R.O., and the work is part of the survey being carried out by the C.S.I.R.O. and other institutions on the chemical constituents of Australian flora.

† Physiology Department, University of Queensland.

‡ Organic Chemistry Department, University of Sydney.



Found: C, 50.3; H, 7.7; N, 7.4; I, 34.0%.

Calculated for  $C_{16}H_{26}N_2I$ : C, 51.0; H, 7.7; N, 7.4; I, 33.8%.

(iii) *d-Sparteine Dihydriodide*

The alkaloid was dissolved in 5% sulphuric acid and a saturated solution of potassium iodide added. On standing, the dihydriodide precipitated and was recrystallized from alcohol as large colourless prisms, m.p. 255–257.5 °C. (lit.(6), m.p. 255–257.5 °C.).

Found: I, 51.0%.

Calculated for  $C_{16}H_{26}N_2 \cdot 2HI \cdot H_2O$ : I, 51.0%.

(iv) *d-Sparteine Perchlorate*

Perchloric acid was added to an ethereal solution of the alkaloid and the precipitated salt recrystallized from alcohol as colourless cubes, m.p. 171.5–172.5 °C. (lit.(3), m.p. 171–172 °C.)

Found: N, 8.2%.

Calculated for  $C_{16}H_{26}N_2 \cdot HClO_4$ : N, 8.4%.

A pK determination by potentiometric titration using a glass electrode showed that there were two basic centres:

$$pK_{a1} = 11.4 \pm 0.1 \quad \text{conc.} = \frac{M}{100} \quad \text{in alcohol water (1:1) at 20 °C.}$$

$$pK_{a2} = 3.96 \pm 0.05 \quad \text{conc.} = \frac{M}{100} \quad \text{in alcohol water (1:1) at 20 °C.}$$

Kolthoff(7) found values for *d*-sparteine in water at 15 °C. to be  $pK_{a1} = 12.0$ ,  $pK_{a2} = 4.89$ .

(b) *Hovea longifolia R.Br.*

This material was collected between Kenmore and Moggill Ferry in October 1947, and after oven-drying worked up in the same way as above. The yield of *d*-sparteine varied between 1.1 and 1.4%, the higher yield being obtained from young plants. The derivatives prepared as above agreed in all respects. A small amount of amorphous material remained after steam distillation which gave strong alkaloid tests but it was insufficient for further examination.

### III. ACKNOWLEDGMENTS

The authors wish to thank Mr. G. K. Hughes, Chemistry Department, University of Sydney, and Mr. H. J. Hines, Physiology Department, University of Queensland, for their interest and advice. The authors are also indebted to Miss J. E. Fildes, microanalyst, University of Sydney, for the semi-microanalyses, to Miss Pam Rolfe for technical assistance, and to Mr. J. N. Phillips for the pK determinations.

### IV. REFERENCES

- (1) WEBB, L. J.—*Coun. Sci. Industr. Res. Aust. Bull.* No. 232 (1948).
- (2) HENRY, T. A.—“*The Plant Alkaloids*,” 3rd Ed. (Blakiston Book Co.: Philadelphia, 1939.)
- (3) ING, H. R.—*J. Chem. Soc.* **1935**: 1053 (1935).
- (4) MARION, L., and OUELLET, J.—*J. Amer. Chem. Soc.* **70**: 691 (1948).
- (5) CLEMO, G. R., RAFFER, R., and TENNISWOOD, C. R. S.—*J. Chem. Soc.* **1931**: 429 (1931).
- (6) ORÉCHOV, A., RABINOWITCH, M., and KANAWALOWA, R. *Ber. dtsch. chem. Ges.* **66**: 612 (1933).
- (7) KOLTHOFF, I. M.—*Biochem. Z.* **162**: 289 (1925).



# ALKALOIDS OF THE AUSTRALIAN RUTACEAE:

## *EVODIA XANTHOXYLOIDES*\*

### I. EVOXANTHINE

By G. K. HUGHES† and K. G. NEILL†

[Manuscript received March 18, 1949]

#### Summary

From the bark and leaves of *E. xanthoxyloides* four alkaloids, melicopidine, *evoxanthine*, *xanthevodine*, and kokusaginine, were isolated. The structure of *evoxanthine* was shown to be 2,3-methylenedioxy-4-methoxy-10-methylacridone.

### I. INTRODUCTION

There are about 120 species of *Evodia* distributed throughout tropical and subtropical Eastern Asia, Polynesia, Australia, and Madagascar. *E. xanthoxyloides* F.v.M. is a small tree found in the rain-forest of north Queensland, most commonly in the Atherton Tableland and the Innisfail-Tully districts. The genus is placed by Engler and Prantl(1) in the family Rutaceae, subfamily Rutoideae, tribe Xanthoxyleae, and subtribe Evodiinae, which also includes known alkaloid-bearing genera such as *Xanthoxylum*, *Fagara*, and *Orixa*. From a methanolic extract of the bright yellow bark three alkaloids have been isolated, melicopidine(2), kokusaginine(3), and the previously unknown *evoxanthine*,  $C_{16}H_{13}O_4N$ . The leaves contain both *evoxanthine* and melicopidine along with another new alkaloid *xanthevodine*,  $C_{17}H_{15}O_5N$ .

Kokusaginine was originally isolated from *Orixa japonica* Thunb., and preliminary investigation suggested it to be a furano-quinoline derivative but the positions of the substituent groups were not determined. Work is in progress in this Laboratory on this problem and it is hoped to report in a future communication.

*Evoxanthine* is a weakly basic yellow solid which shows an intense blue-violet fluorescence in alcoholic solution. It contains one methoxy group and gives a positive gallic acid-sulphuric acid test for a methylenedioxy group. The remaining oxygen atom showed no carbonyl or phenolic properties.

On nitric acid oxidation it gave 1-methyl-4-quinolone-3-carboxylic acid (cf. Price 4), and was thus structurally related to melicopidine and some of the other alkaloids isolated from *Melicope fareana* F.v.M., that is, it was a substituted

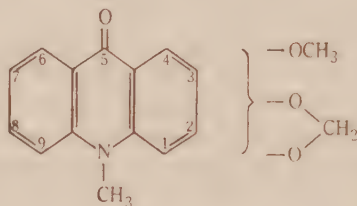
\* This material was supplied by Mr. L. J. Webb, C.S.I.R.O., and is part of the survey being carried out by C.S.I.R.O. in collaboration with other Australian organizations.

† Organic Chemistry Department, University of Sydney.

10-methylacridone with a methoxyl and methylenedioxy in one ring of the nucleus.

The relative positions of the methoxy and methylenedioxy groups were determined as follows.

Evoxanthine on heating with alcoholic hydrochloric acid gave an orange crystalline solid,  $C_{15}H_{11}O_4N$ , which contained no methoxyl group and will be termed *norevioxanthine*. It is an extremely weak base soluble only in hot concentrated hydrochloric acid and insoluble in sodium hydroxide. However, it gave a monoacetate with difficulty and was remethylated to evoxanthine.



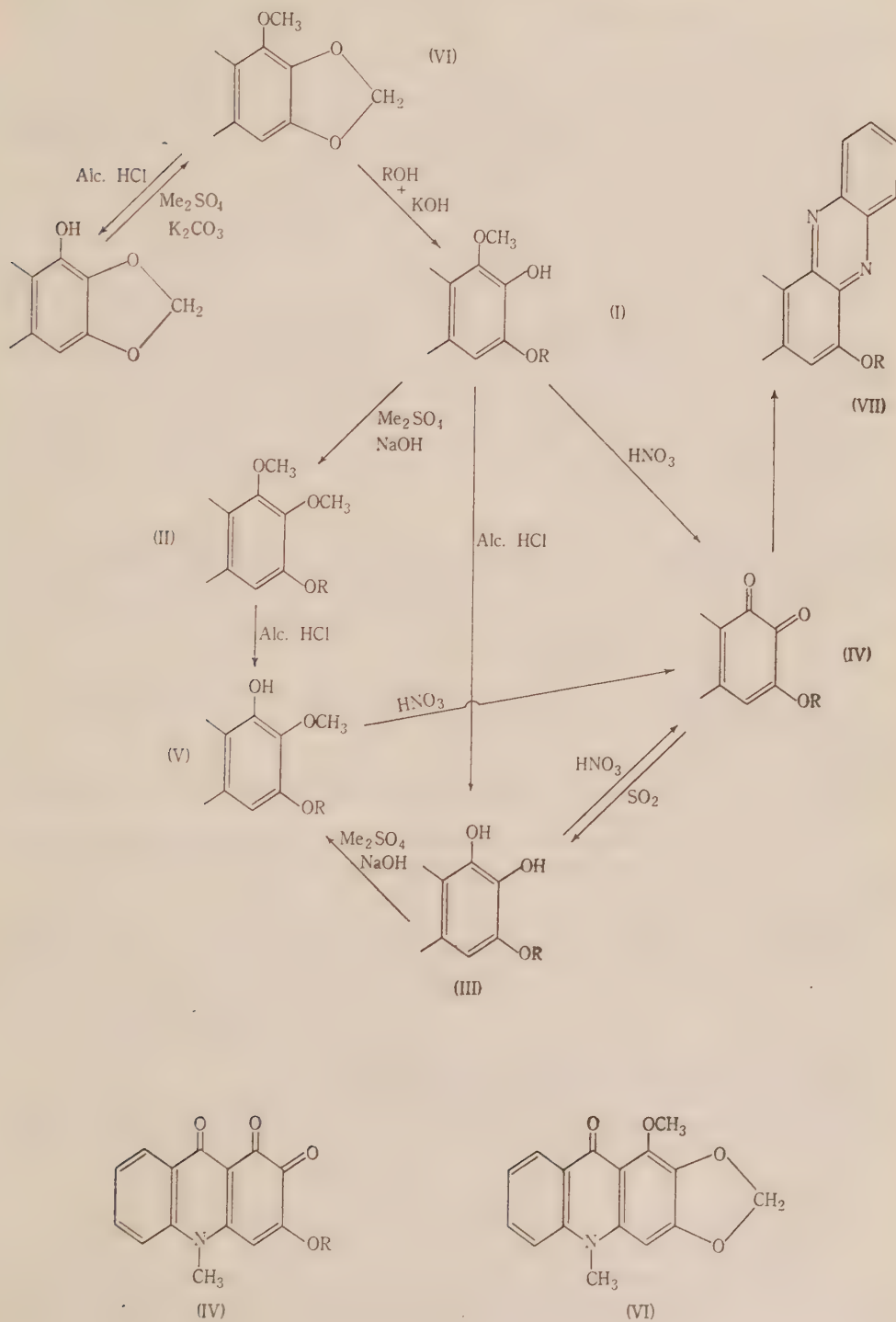
On heating for some hours with methanolic potassium hydroxide, evoxanthine gave a yellow phenol,  $C_{16}H_{15}O_4N$  (Ia, R = Me), which was readily methylated to  $C_{17}H_{17}O_4N$  (IIa, R = Me). If ethanolic potassium hydroxide was used another phenol,  $C_{17}H_{17}O_4N$  (Ib, R = Et), was obtained; this was further methylated to  $C_{18}H_{19}O_4N$  (IIb, R = Et). Neither of the above phenols gave a positive gallic acid-sulphuric acid test and therefore must be formed by the rupture of the methylenedioxy ring (cf. melicopine and melicopidine).

On treatment with alcoholic hydrochloric acid these phenols were demethylated to give  $C_{15}H_{13}O_4N$  (IIIa), and  $C_{16}H_{15}O_4N$  (IIIb), respectively, both of which crystallize as yellow-orange photosensitive needles and contain one alkoxy group and two hydroxyl groups.

On treatment with cold concentrated nitric acid IIIa and IIIb gave dark red crystalline solids,  $C_{15}H_{11}O_4N$  (IVa) and  $C_{16}H_{13}O_4N$  (IVb), respectively, which were readily reduced back by sulphur dioxide to the dihydroxy compounds. This easily reversible oxidation-reduction involving two hydrogen atoms, together with yellow and red colours of the compounds, was characteristic of a quinol-quinone system, and as the latter compounds reacted with *o*-phenylenediamine to give monoalkoxyphenazines (VII) they were 1,2-quinones.

Both quinols (IIIa and IIIb) yielded monomethyl ethers (Va and Vb), isomeric with Ia and Ib but could not be further methylated under the most vigorous conditions. The same difficulty was met in the diacetylation, a small yield of IIIb diacetate being obtained after long heating with acetic anhydride and sodium acetate.

Crow and Price(5) have shown that the easily demethylated methoxyl group in three alkaloids of this series is attached to the 4 position. Now, as the alkoxy group formed by the rupture of the methylenedioxy ring is retained during the phenazine formation the *ortho*-quinones cannot arise from the methylenedioxy ring alone and must therefore be 3,4-quinones (IV, a, R = Me; b, R = Et). From this it follows that evoxanthine must have structure VI.



## II. EXPERIMENTAL

All melting points are uncorrected unless otherwise stated.

(i) *Extraction of Bark*

The dried milled bark was exhausted by cold percolation with methanol, the solvent removed, the residual dark oil refluxed with chloroform, and the whole allowed to stand in the refrigerator for some days. The chloroform was decanted from a thick sugary syrup, concentrated, and on standing deposited almost pure evoxanthine.

The mother liquor was extracted with sodium hydroxide to remove acidic constituents which were discarded, and the chloroform removed. The black residue was treated with boiling acetone, and again allowed to stand, when a further insoluble deposit separated. To the cold supernatant liquor concentrated hydrochloric acid was added until no further precipitation occurred, the hydrochlorides collected, washed with ice cold acetone, and then decomposed with ammonia. The crude alkaloid mixture (evoxanthine, kokusagine, and a small amount of melicopidine) was dried, dissolved in chloroform, and on cooling, the bulk of the evoxanthine crystallized. The mother liquors were evaporated to dryness, the residue refluxed with alcoholic hydrochloric acid (5% concentrated hydrochloric acid by volume) for 3 hours, concentrated, cooled, and the mixture of *norevioxanthine*, *normelicopidine*, and kokusagine hydrochloride collected. This was extracted with hot 2% hydrochloric acid to dissolve the kokusagine hydrochloride which crystallized on cooling and was then converted to the free base with ammonia.

The *normelicopidine* was removed from the *norevioxanthine* by boiling with chloroform in which the latter is insoluble.

From the original acid-acetone filtrate the main yield of *normelicopidine* plus further small quantities of kokusagine and *norevioxanthine* were recovered after removing most of the solvent and separating as above.

The yield of evoxanthine (calculated from pure evoxanthine and *norevioxanthine*) varied from tree to tree and was approximately 1%. The yield of *normelicopidine* was 0.8%, and the yield of kokusagine 0.1–0.2%.

Pure melicopidine could be isolated with difficulty from the above acetone solution by neutralization and repeated recrystallization from methanol.

(ii) *Extraction of Leaves*

The extraction and separation were similar to those carried out on the bark: xanthevodine was soluble in 2% hydrochloric acid.

Further work is being carried out on the extraction of the leaves and in a further communication on the structure of xanthevodine the method of isolation will be discussed in more detail.

(iii) *Evioxanthine*

Evioxanthine crystallizes in yellow needles from benzene, toluene, ethyl acetate-light petroleum, and chloroform, or in cubes from absolute alcohol, m.p. 217–218 °C. It crystallizes with solvent of crystallization from chloroform but loses this on exposure to air. It is a weak base having a  $pK_a = 2.6 \pm 0.2$  at 20 °C. ( $M_{200}$  in 50% ethanol) and is soluble in warm 2% hydrochloric acid from which the hydrochloride crystallizes on cooling.

Found: C, 67.8; H, 4.5; N, 5.0;  $CH_3O$ , 10.9;  $NCH_3$ , 5.3%.

Calculated for  $C_{16}H_{15}O_4N$ : C, 67.8; H, 4.6; N, 5.0;  $CH_3O$ , 10.9;  $NCH_3$ , 5.3% (calculated as methyl).

(iv) *Melicopidine*

Melicopidine crystallized in glistening yellow needles from methanol, m.p. and mixed m.p., with authentic specimen from *M. fareana*, 121–122 °C. (corr.).

*nor*Melicopidine crystallized as orange needles from chloroform or as long glistening red needles with a coppery lustre from toluene or benzene, m.p. 208.5–209 °C., mixed m.p. with an authentic specimen, 208–209 °C.



(v) *Kokusaginine*

Kokusaginine crystallized as colourless needles from alcohol or as prisms from benzene m.p. 170–171 °C. Terasaka(3) gives m.p. 171 °C.

Found: C, 64.3; H, 5.2; N, 5.4;  $\text{CH}_3\text{O}$ , 33.4%.

Calculated for  $\text{C}_{14}\text{H}_{13}\text{O}_4\text{N}$ : C, 64.9; H, 5.1; N, 5.4;  $3 \times \text{CH}_3\text{O}$ , 35.9%.

(vi) *Kokusaginine Picrate*

The picrate was precipitated in alcohol and crystallized from acetone as fine golden needles m.p. 218.5–219.5 °C. Terasaka(3) gives m.p. 218 °C.

Found: N, 11.1%.

Calculated for  $\text{C}_{14}\text{H}_{13}\text{O}_4\text{N} \cdot \text{C}_6\text{H}_3\text{O}_7\text{N}_3$ : N, 11.5%.

(vii) *norEvoxanthine*

Evoxanthine (5 g.) dissolved in alcohol (300 cc.) and concentrated hydrochloric acid (15 cc.) were warmed on the water-bath, and after an hour orange needles separated. After 2 hours the "nor" compound was collected and crystallized from pyridine or dioxan as orange needles, m.p. 274–275 °C. *norEvoxanthine* is a very weak base and is soluble only in warm concentrated hydrochloric acid. The yield is quantitative.

Found: C, 66.8; H, 4.2; N, 5.1%;  $\text{CH}_3\text{O}$ , nil.

Calculated for  $\text{C}_{15}\text{H}_{11}\text{O}_4$ : C, 66.9; H, 4.1; N, 5.2%;  $\text{CH}_3\text{O}$ , nil.

(viii) *Acetyl norEvoxanthine*

*norEvoxanthine* (0.5 g.) was refluxed with acetic anhydride (20 cc.) and anhydrous sodium acetate (0.1–0.2 g.) for 10 hours, cooled, filtered, and crystallized as yellow needles from ethyl acetate-light petroleum, m.p. 240–242 °C. (soften at 235 °C.).

Found: C, 65.7; H, 4.1; N, 4.5%.

Calculated for  $\text{C}_{15}\text{H}_{11}\text{O}_4\text{N}(\text{C}_2\text{H}_2\text{O})$ : C, 65.6; H, 4.2; N, 4.5%.

(ix) *Methylation of norEvoxanthine*

*norEvoxanthine* (0.75 g.) was refluxed for 24 hours with excess dimethyl sulphate and anhydrous potassium carbonate in acetone. The potassium salts were removed and the *evoxanthine* recovered through its hydrochloride and after recrystallization from toluene had m.p. and mixed m.p., with a natural specimen, 217–218 °C.

(x) *1-Methyl-4-Quinolone-3-Carboxylic Acid*

To *evoxanthine* (1 g.) was added concentrated nitric acid (5–10 cc.); as soon as frothing subsided somewhat the nitric acid was boiled off as quickly as possible, the residual brownish gum was dissolved immediately in boiling glacial acetic acid, and crystallized on standing as cream needles, m.p. and mixed m.p. with an authentic specimen prepared from the melicope alkaloids, 296–297 °C. (corr.).

(xi) *2,4-Dimethoxy-3-Hydroxy-10-Methylacridone (Ia)*

*Evoxanthine* (10 g.) was refluxed for 24 hours with methanol (600 cc.) and potassium hydroxide (50 g.), the solvent removed, water added, and then extracted with chloroform. On evaporation, this yielded a mixture of unchanged *evoxanthine* and *norevoxanthine*, identified by m.p. and mixed m.p. The aqueous alkaline solution was then saturated with carbon dioxide and the precipitated phenol (*Ia*) collected and crystallized from toluene or ethyl acetate-light petroleum in glistening yellow needles which are photosensitive, m.p. 226–227 °C.

Found: C, 67.2; H, 5.1; N, 5.1;  $\text{CH}_3\text{O}$ , 21.5%.

Calculated for  $\text{C}_{16}\text{H}_{15}\text{O}_4\text{N}$ : C, 67.3; H, 5.3; N, 4.9;  $2 \times \text{CH}_3\text{O}$ , 21.8%.

(xii) *Acetyl Derivative of Ia*

The phenol was acetylated with acetic anhydride and pyridine in the usual way and the product recrystallized from aqueous alcohol in off-white needles, m.p. 169.5–171 °C.

Found: C, 65.7; H, 5.3; N, 4.9%.

Calculated for  $C_{16}H_{15}O_4N(C_2H_5O)$ : C, 66.0; H, 5.2; N, 4.3%.

(xiii) *2,3,4-Trimethoxy-10-Methylacridone (IIa)*

The phenol (Ia) was dissolved in dilute potassium hydroxide containing sodium hydro-sulphite, excess dimethyl sulphate added, and shaken. After a short time the mixture became warm, and the methyl ether which separated was filtered, washed, dried, and crystallized from benzene-light petroleum in pale yellow plates, m.p. 168.5–170 °C.

Found: C, 68.3; H, 5.4; N, 4.8;  $CH_3O$ , 29.8%.

Calculated for  $C_{17}H_{17}O_4N$ : C, 68.3; H, 5.7; N, 4.7;  $3 \times CH_3O$ , 31.1%.

(xiv) *2-Ethoxy-3-Hydroxy-4-Methoxy-10-Methylacridone (Ib)*

Evioxanthine (5 g.) was refluxed for 8 hours with absolute ethanol (500 cc.) and potassium hydroxide (25 g.). The phenol (Ib), isolated in a similar manner to Ia, crystallized from benzene in glistening flat yellow prisms, m.p. 199–201 °C. (sinters at 195 °C.).

Found: C, 68.2; H, 5.6; N, 4.7;  $NCH_3$ , 4.6;  $CH_3O$ , 19.0%.

Calculated for  $C_{17}H_{17}O_4N$ : C, 68.3; H, 5.7; N, 4.7;  $NCH_3$ , 5.0;  $2 \times CH_3O$ , 20.7%.

The alkoxylys were calculated as two methoxyls.

(xv) *Acetyl Derivative of Ib*

The phenol was acetylated in the same manner as phenol (Ia), and crystallized from aqueous alcohol in glistening colourless needles, m.p. 209–210 °C.

Found: C, 67.0; H, 5.6; N, 4.3;  $CH_3CO$ , 11.6%.

Calculated for  $C_{17}H_{17}O_4N(C_2H_5O)$ : C, 66.9; H, 5.6; N, 4.1;  $CH_3CO$ , 12.6%.

(xvi) *2-Ethoxy-3,4-Dimethoxy-10-Methylacridone (IIb)*

The product obtained from the methylation of Ib crystallized from benzene-light petroleum in long, yellow needles, m.p. 141–142 °C.

Found: C, 68.6; H, 6.1; N, 4.6;  $CH_3O$ , 28.8%.

Calculated for  $C_{18}H_{19}O_4N$ : C, 69.0; H, 6.1; N, 4.5;  $3 \times CH_3O$ , 29.7%.

(xvii) *2-Methoxy-3,4-Dihydroxy-10-Methylacridone (IIIa)*

(a) *From Ia.*—The phenol (0.5 g.) was heated on a water-bath with alcohol (30 cc.) and concentrated hydrochloric acid (30 cc.) for 4 hours and the "nor" compound crystallized in glistening golden needles from methanol containing sulphur dioxide or from toluene, m.p. 241–242 °C.

(b) *From IVa.*—The quinone (0.5 g.) was suspended in dilute sodium carbonate solution and sodium hydrosulphite added. On stirring the red quinone was rapidly reduced to the yellow quinol which crystallized from toluene or methanol containing sulphur dioxide, m.p. and mixed m.p. with product from above, 241–242 °C.

Found: C, 66.1; H, 5.1; N, 5.2;  $CH_3O$ , 10.5%.

Calculated for  $C_{15}H_{13}O_4N$ : C, 66.3; H, 5.0; N, 5.2;  $CH_3O$ , 11.4%.

(xviii) *Monoacetyl Derivative of IIIa*

The quinol (IIIa) was dissolved in hot acetic anhydride, cooled, and the acetyl derivative filtered, and crystallized from ethyl acetate-light petroleum in yellow plates, m.p. 255–257 °C.

Found: C, 64.8; H, 4.8; N, 4.5%.

Calculated for  $C_{16}H_{15}O_4N(C_2H_5O)$ : C, 65.2; H, 4.8; N, 4.5%.

(xix) *Diacetyl Derivative of IIIa*

All attempts to prepare this compound either from the quinone (IVa), the quinol (IIIa), or its monoacetate by refluxing with acetic anhydride and anhydrous sodium acetate yielded the monoacetate, m.p. and mixed m.p. 255–257 °C. There was evidence of the presence of a small amount of the diacetate as solutions of the crude product gave a blue fluorescence characteristic of the diacetate of IIIb and not given by the pure monoacetate or the starting materials.

(xx) *2,3-Dimethoxy-4-Hydroxy-10-Methylacridone (Va)*

(a) *From IIIa*.—The quinol (IIIa), (0.5 g.) was dissolved in 10% sodium hydroxide solution (50 cc.) containing sodium hydrosulphite (0.2 g.), dimethyl sulphate (2 cc.) added, and shaken vigorously for several minutes; the red colour of the solution gradually faded and the methyl ether separated as a yellow solid and crystallized from alcohol in long golden needles, m.p. 176–177.5 °C.

(b) *From IIa*.—The ether (0.05 g.) was refluxed with alcoholic hydrochloric acid (8 cc.) for 2 hours, the “nor” compound was identical with the ether from above, m.p. and mixed m.p., 176–177.5 °C.

Found: C, 66.4; H, 5.6; N, 5.2; CH<sub>3</sub>O, 20.7%.

Calculated for C<sub>16</sub>H<sub>15</sub>O<sub>4</sub>N: C, 67.4; H, 5.3; N, 4.9; 2 × CH<sub>3</sub>O, 21.8%.

(xxi) *2-Ethoxy-3,4-Dihydroxy-10-Methylacridone (IIIb)*

(a) *From Ib*.—Prepared by the same method as IIIa and crystallized from alcohol in golden needles which were photosensitive, m.p. 258.5–259 °C. (immersed at 245 °C.).

(b) *From IVb*.—The quinone (IVb) was suspended in water and sulphur dioxide bubbled through it; the quinol did not depress the m.p. of product from above.

Found: C, 67.2; H, 5.4; N, 4.9; CH<sub>3</sub>O, 9.7%.

Calculated for C<sub>16</sub>H<sub>15</sub>O<sub>4</sub>N: C, 67.4; H, 5.3; N, 4.9; 1 × CH<sub>3</sub>O, 10.9%.

(xxii) *Monoacetate of IIIb*

The quinol (IIIb) was acetylated by the same method as for IIIa and crystallized from dioxan in yellow, flat needles, m.p. 255–256 °C.

Found: C, 65.8; H, 5.5; N, 4.2%.

Calculated for C<sub>18</sub>H<sub>15</sub>O<sub>4</sub>N(C<sub>2</sub>H<sub>2</sub>O): C, 66.0; H, 5.2; N, 4.3%.

(xxiii) *Diacetate of IIIb*

The monoacetate was refluxed for 8 hours in acetic anhydride with anhydrous sodium acetate, cooled, filtered, and crystallized from ethyl acetate-light petroleum in colourless needles, m.p. 229–230 °C.

Found: C, 62.3; H, 5.3; N, 4.1%.

Calculated for C<sub>18</sub>H<sub>15</sub>O<sub>4</sub>N(C<sub>2</sub>H<sub>2</sub>O)<sub>2</sub>: C, 65.0; H, 5.2; N, 3.8%.

(xxiv) *2-Ethoxy-3-Methoxy-4-Hydroxy-10-Methylacridone (Vb)*

(a) *From IIIb*.—By the same method as for Va, crystallized from alcohol in long golden needles, m.p. 193.5–194 °C.

(b) *From IIb*.—By the same method as for Va; the product did not depress m.p. of compound obtained from method above.

Found: C, 68.2; H, 5.7; N, 4.8; CH<sub>3</sub>O, 21.2%.

Calculated for C<sub>17</sub>H<sub>17</sub>O<sub>4</sub>N: C, 68.3; H, 5.7; N, 4.7; 2 × CH<sub>3</sub>O, 20.7%.

(xxv) *2-Ethoxy-3-Methoxy-4-Acetoxy-10-Methylacridone*

Vb was refluxed with acetic anhydride and anhydrous sodium acetate for 8 hours, cooled, collected, and crystallized from ethyl acetate-light petroleum in long yellow needles, m.p. 210.5–211 °C.

Found: C, 66.1; H, 5.8; N, 4.3%.

Calculated for C<sub>19</sub>H<sub>19</sub>O<sub>5</sub>N: C, 66.9; H, 5.6; N, 4.1%.

(xxvi) *2,3-Diethoxy-4-Hydroxy-10-Methylacridone*

On treatment with diethyl sulphate and caustic soda, IIIb yielded the yellow ether which crystallized in golden needles from alcohol or benzene-light petroleum, m.p. 173 °C.

Found: C, 67.5; H, 5.9; N, 4.4; C<sub>2</sub>H<sub>5</sub>O, 26.6%.

Calculated for C<sub>18</sub>H<sub>18</sub>O<sub>4</sub>N: C, 69.0; H, 6.1; N, 4.5; 2 × C<sub>2</sub>H<sub>5</sub>O, 28.8%.



(xxvii) *2-Methoxy-10-Methylacridone 3,4-Quinone (IVa)*

Ia, IIIa, or Va (0.5 g.) was mixed into a paste with a little water, cooled in ice, cold concentrated nitric acid (10 cc.) added, and stirred until the solid had dissolved to give a bright red solution, allowed to stand for 2 to 3 minutes, and poured onto a paste of excess sodium bicarbonate. The quinone was filtered, washed, and crystallized from a large volume of slightly acid hot water, as dark red needles, m.p. 279–280 °C. (decomp.). If about half the quantity of nitric acid was used the yellow quinol crystallized from the nitric acid and on addition of more nitric acid dissolved. The quinone forms a very stable hydrate which could not be dehydrated by drying in a pistol at 118 °C. (2 mm.). It was dehydrated by mixing with sodium dried toluene and distillation of the water-toluene mixture, repetition of this several times yielding a product which analysed for the anhydrous compound.

Found: C, 61.5; H, 4.6; N, 4.8;  $\text{CH}_3\text{O}$ , 10.6%.

Calculated for  $\text{C}_{15}\text{H}_{11}\text{O}_4\text{N}\cdot\text{H}_2\text{O}$ : C, 62.7; H, 4.5; N, 4.9;  $\text{CH}_3\text{O}$ , 10.8%.

(a) *Dehydrated*

Found: C, 66.1; H, 4.2; N, 5.3;  $\text{CH}_3\text{O}$ , 10.8%.

Calculated for  $\text{C}_{15}\text{H}_{11}\text{O}_4\text{N}$ : C, 66.9; H, 4.1; N, 5.2;  $1 \times \text{CH}_3\text{O}$ , 11.2%.

(xxviii) *2-Ethoxy-10-Methylacridone 3,4-Quinone (IVb)*

The quinone was obtained by treating either Ib, IIIb, or Vb with nitric acid as for IVa; it crystallized from hot slightly acid water in dark red needles, m.p. 261 °C. (decomp.).

Found: C, 67.5; H, 4.5; N, 4.9;  $\text{C}_2\text{H}_5\text{O}$ , 15.3%.

Calculated for  $\text{C}_{16}\text{H}_{13}\text{O}_4\text{N}$ : C, 67.8; H, 4.6; N, 5.0;  $\text{C}_2\text{H}_5\text{O}$ , 15.9%.

(xxix) *Action of o-Phenylenediamine on Quinone (IVa)*

The quinone (IVa), (0.5 g.) was dissolved in warm 10% hydrochloric acid and o-phenylenediamine hydrochloride (0.5 g.) added, the solution immediately darkened and was heated for 5–10 minutes, cooled, and basified with ammonia. The brownish coloured precipitate crystallized from alcohol in fine golden needles, m.p. 285–287 °C.

Found: C, 69.2; H, 4.8; N, 11.3;  $\text{CH}_3\text{O}$ , 8.4; O, 13.6%.

Calculated for  $\text{C}_{21}\text{H}_{15}\text{O}_2\text{N}_3\cdot\text{H}_2\text{O}$ : C, 70.1; H, 4.8; N, 11.7;  $\text{CH}_3\text{O}$ , 8.6; O, 13.4%.

(xxx) *Action of o-Phenylenediamine on Quinone (IVb)*

The quinone (IVb), (0.5 g.) was treated by the same method as the quinone (IVa) and the product crystallized from alcohol in fine golden needles, m.p. 304–305 °C.

Found: N, 11.7;  $\text{C}_2\text{H}_5\text{O}$ , 11.8%.

Calculated for  $\text{C}_{22}\text{H}_{17}\text{O}_2\text{N}_3$ : N, 11.8;  $\text{C}_2\text{H}_5\text{O}$ , 12.7%.

## III. ACKNOWLEDGMENTS

The authors wish to thank Mr. L. J. Webb, C.S.I.R.O., for the collection of the bark and the botanical description, Dr. J. R. Price, C.S.I.R.O., for his interest, advice, and advance information of the melicope alkaloids, Mr. W. D. Crow, C.S.I.R.O., for some preliminary investigation, Miss J. E. Fildes, micro-analyst, University of Sydney, and Mr. R. B. Bradbury, C.S.I.R.O., for the analyses, and Mr. J. N. Phillips for the pK determinations.

## IV. REFERENCES

- (1) ENGLER, A., and PRANTL, K. "Die natürlichen Pflanzenfamilien." Vol. 19a. (Wilhelm Engelmann: Leipzig, 1931.)
- (2) CROW, W. D., and PRICE, J. R.—*Aust. J. Sci. Res. A* **2**: 255 (1949).
- (3) TERASAKI, M.—*J. Pharm. Soc. Japan* **53**: 219 (1933); *Chem. Zbl.* **105**: 405 (1934); *Chem. Abstr.* **29**: 7336 (1935).
- (4) PRICE, J. R.—*Aust. J. Sci. Res. A* **2**: 272 (1949).
- (5) CROW, W. D., and PRICE, J. R.—*Ibid.* **2**: 282 (1949).



# DIFFUSION IN A FIELD OF HOMOGENEOUS TURBULENCE

## I. EULERIAN ANALYSIS

By G. K. BATCHELOR\*

[*Manuscript received July 26, 1949*]

### *Summary*

This paper considers the mechanism of diffusion in the simple case of a homogeneous field of turbulence. It is shown to be useful to distinguish between diffusion from fixed and moving centres, and only the former is considered here. The diffusion of a cloud of marked fluid particles about the average position of their centre is known when the statistical behaviour of a single fluid particle is known. Theory shows that the dispersion of a fluid particle about its initial position increases first as the square of the time of flight,  $t$ , then more slowly, and eventually increases linearly in  $t$ . Several different experiments have shown that the probability distribution of the displacement of a fluid particle is normal for *all* values of the time of flight. As a consequence of these two facts, it is possible to represent the diffusion by a differential equation of the heat-conduction type, with a diffusion coefficient which initially increases with  $t$  and eventually becomes constant. Some consequences of the analysis are presented. Part II of this paper will discuss the more important case of diffusion about a moving centre.

## I. INTRODUCTION

The problem of diffusion in turbulent motion lies somewhat apart from the topics which have received most attention in the literature on turbulence, but is none the less of intrinsic interest and of practical importance. Diffusion can be regarded as almost wholly a kinematic problem, depending only indirectly on the dynamics of turbulent motion. We rely on analysis of the type given in the literature(1) for information about the statistical properties of the turbulence, and a discussion of turbulent diffusion need be concerned—in the first place at any rate—only with the definition and description of the diffusion in terms of the characteristics of the turbulence. For instance, an interesting problem is to determine the rate of increase of the width of the heat wake with distance from a heated obstacle in a turbulent stream. The first enquiry should be, how is this rate of spreading described in terms of the statistical characteristics of the turbulence? ; and the second enquiry concerns the determination, either from given data about the field of turbulence or from *a priori* dynamical deductions, of these relevant characteristics of the turbulence. The first kind of enquiry is the one to be discussed in this paper.

Unfortunately, most of the statistical characteristics of turbulence found to be relevant to the description of diffusion are both unmeasurable (with present equipment and techniques) and outside the scope of prediction from the existing dynamical

\*Trinity College, Cambridge.

theory. Probably for this reason, turbulent diffusion has not been studied by many authors. Taylor's pioneering paper(2) published 28 years ago is a notable exception, and papers written by Richardson(3) and Sutton(4) from a meteorological point of view should also be mentioned. We must, therefore, as a beginning, choose the simplest possible conditions, even though these conditions may seem to bear little resemblance to those encountered in the important practical problems. The field of turbulent motion will be assumed herein to be statistically homogeneous, but no restrictions are placed upon the directional characteristics. The turbulence will also be assumed to be statistically steady in time; this will be approximately valid even in a decaying field provided the time interval correlations occurring in the analysis are close to zero for intervals which are small compared with a time in which

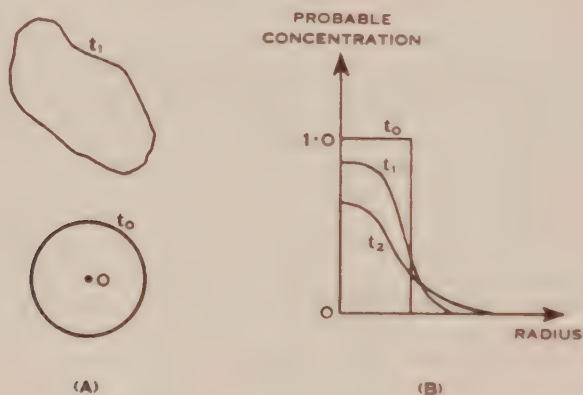


Fig. 1

the turbulence characteristics change appreciably. Mean values will be regarded as stochastic averages, i.e. as averages over a large number of independent experiments or trials under similar conditions, the difference between any two trials being a consequence of the randomness inherent in the turbulent motion. The effect of molecular diffusion will be ignored.

We consider the diffusion of some transportable property or quality which at some initial instant is possessed by the fluid particles within a given marked volume. The diffusing quality is assumed not to confer any new dynamical properties on the fluid particles which are transporting it. The nature of the quality concerned is not important. It may be either of the discrete kind—for example the marked fluid particles might be composed of another kind of molecule—or it may be capable of continuous variation like temperature. The essential feature of purely turbulent diffusion is that the fluid particles retain their initial properties throughout the history of the motion, but the spatial distribution of the particles changes. If the marked fluid particles (those carrying the property concerned) fill a volume,  $V$ , the total volume of the marked particles at all subsequent times will be  $V$  and connected volumes will remain connected, although the shape of the enclosing surface will change considerably with time.

Two distinct kinds of problem appear to arise in a consideration of turbulent diffusion, and these call for two different kinds of analysis. In what might be called an Eulerian type of analysis we enquire into the *probable* concentration of the diffusing property at *fixed* points of space at different instants. In a Lagrangian type of analysis we enquire into the probable distribution of the diffusing property with respect to points which *follow the fluid motion*. A simple example will make clear the difference. Consider the diffusion of some property possessed with unit intensity by all the fluid particles which initially (time  $t_0$ ) lie within a sphere centred at a point  $O$ . At a later instant,  $t$ , the cloud will be distorted and displaced, a typical result for any one trial being as shown in Figure 1(A). In an Eulerian type of analysis of this problem, the aim is to determine the probability  $P(x, y, z, t)$  that the point  $(x, y, z)$  lies within the volume of the marked fluid at some later time,  $t$ ; in a practical experiment  $P(x, y, z, t)$  would be identical with the mean concentration (over a large number of trials) of the diffusing property at the point  $(x, y, z)$  and time  $t$ . If the turbulence has statistical spherical symmetry, contours of equal values of  $P$  at any instant are spheres centred at  $O$ , and the radial distribution of  $P$  would be roughly as in Figure 1(B). On the other hand, in a Lagrangian type of analysis we should attempt to determine the average geometrical characteristics of the moving volume of marked fluid particles; the mean dispersion of the cloud about its (moving) centre of mass is one such characteristic. The Eulerian specification of the diffusion may sometimes be recovered from the Lagrangian by an appropriate spatial integration, but not *vice versa*.

It is clear that the usefulness of either type of analysis in any particular problem will depend largely on the ratio of the scale of the turbulence to the linear dimensions of the configuration of marked particles which is being diffused. When this ratio is small, the boundary of the cloud of marked particles undergoes small-scale distortion but the centre of mass of the cloud is approximately stationary. Under these circumstances an Eulerian type of analysis will supply all the information likely to be of interest. On the other hand, when the scale of turbulence is much larger than the cloud dimensions, the relative motion of parts of the cloud will be small compared with the translation of the cloud as a whole. Since the distortion, and not the translation, will normally be of interest, it will be necessary to *follow* the motion of the cloud and a Lagrangian type of analysis is required. In the important practical case of a fixed continuously emitting source of diffusible substance, interest lies in diffusion relative to the (fixed) position of the source, and Eulerian analysis is appropriate.

The theoretical importance of the distinction between the two types of analysis lies in the fact that, in general, Lagrangian analysis will involve correlations between the velocities of two different fluid particles at the same or at different instants, since such correlations will specify the relative motion of different parts of the cloud, while Eulerian analysis will involve only correlations between the velocities of a single fluid particle at different instants. In this paper in which simplicity of conditions is at a premium, we confine attention to the Eulerian type of analysis. The Lagrangian type of analysis will be considered in a later paper.



## II. THE DISPERSION OF A SINGLE FLUID PARTICLE

Since the Eulerian type of analysis is concerned with the probability of finding marked fluid at a given point of space at a given time, we can discuss in isolation each particle of the original volume of marked fluid. Each of these particles is displaced by the turbulent motion, and the probability of finding it in a certain position at a later time is not affected by the existence of other marked fluid particles. Thus, knowing the statistical characteristics of the motion of a single fluid particle we can build up the Eulerian (but not the Lagrangian) specification of the diffusion of a cloud of marked particles. The opposite conclusion, viz. "a spreading dot will not serve as an element from which more general distributions can be built up" has been stated by Richardson(3), but this is true only of a Lagrangian specification of the diffusion.

Let  $X_i(t)$  be the  $i$ -component of the displacement of a fluid particle over an interval  $t$  seconds in any one trial. The method of relating the mean square (the statistical quantity of most interest) of  $X_i$  to a double velocity correlation has been given by Taylor(2). Here we generalize to the case of three dimensions and steady homogeneous turbulence. Representing a mean value by an overbar, the rate of change of the generalized dispersion  $\overline{X_i(t)X_j(t)}$  at time  $t$  is

$$\begin{aligned}\frac{d\overline{X_iX_j}}{dt} &= \overline{X_iu_j} + \overline{u_iX_j} \\ &= \int_0^t \overline{u_i(t')u_j(t)}dt' + \int_0^t \overline{u_i(t)u_j(t')}dt', \quad \dots\dots\dots(2.1)\end{aligned}$$

where  $u_i(t)$  is the  $i$ -component of the velocity of the fluid particle at time  $t$ .  $\overline{u_i(t')u_j(t)}$  is the correlation between components of velocity of the moving fluid particle at two different instants and is consequently a typical component of a second-order tensor which is a function only of the interval  $t - t'$ . We write

$$\overline{u_i(t')u_j(t)} = S_{ij}(t - t'), \quad \dots\dots\dots(2.2)$$

and thus

$$\frac{d\overline{X_iX_j}}{dt} = \int_0^t [S_{ij}(\xi) + S_{ji}(\xi)]d\xi \quad \dots\dots\dots(2.3)$$

The generalized dispersion of the fluid particle at time  $t$  is therefore a second-order tensor of which a typical component is

$$\begin{aligned}\overline{X_iX_j} &= \int_0^t \int_0^{\tau} [S_{ij}(\xi) + S_{ji}(\xi)]d\xi d\tau \\ &= \int_0^t (t - \xi)[S_{ij}(\xi) + S_{ji}(\xi)]d\xi \quad \dots\dots\dots(2.4)\end{aligned}$$

The form of the tensor function  $S_{ij}(\xi)$  cannot be predicted generally, but in two limiting cases the expression (2.4) simplifies. When  $t$  is so small that the integrand



$S_{ij}(\xi) + S_{ji}(\xi)$  can be represented by its value at  $\xi = 0$ ,

$$\overline{X_i X_j} \approx t^2 \overline{u_i u_j} \quad \dots\dots\dots(2.5)$$

In this case the particle moves with the same velocity throughout the whole of the time  $t$ . At the other extreme it is reasonable to assume that when  $\xi$  is large,  $S_{ij}(\xi)$  and  $S_{ji}(\xi)$  are very small and that

$$\left. \begin{aligned} \int_0^t [S_{ij}(\xi) + S_{ji}(\xi)] d\xi &\rightarrow I^{(0)}_{ij} \\ \int_0^t \xi [S_{ij}(\xi) + S_{ji}(\xi)] d\xi &\rightarrow I^{(1)}_{ij} \end{aligned} \right\}, \dots\dots\dots(2.6)$$

as  $t$  tends to infinity, where  $I^{(0)}_{ij}$  and  $I^{(1)}_{ij}$  are constants. In this case (2.4) becomes

$$\overline{X_i X_j} \rightarrow t I^{(0)}_{ij} - I^{(1)}_{ij} \text{ as } t \rightarrow \infty \quad \dots\dots\dots(2.7)$$

This linear expression for the dispersion corresponds to that known to be valid for molecular diffusion. In this latter case a molecule loses correlation with its initial motion after it has made a few collisions, so that the asymptotic condition is set up almost immediately by comparison with observational intervals of time. On the other hand, in the atmosphere where there is an enormous range of eddy sizes, the asymptotic condition (2.7) may *not* be attained within practical intervals of time.

In a paper devoted to a discussion of the mathematical foundations of the analysis of stationary random functions like  $X_i(t)$  and  $u_i(t)$ , Kampé de Fériet(5) has pointed out that the dispersion can be related to the spectrum of the kinetic energy of the fluid particle. Again generalizing to three dimensions, we define the spectrum tensor of the *velocity of a particle* (as distinct from the corresponding spectrum tensor of the velocity at a point in space; see Batchelor 6) as

$$\Phi_{ij}(n) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S_{ij}(\xi) e^{-i\xi n} d\xi, \quad \dots\dots\dots(2.8)$$

where  $i = \sqrt{-1}$  and  $n/2\pi$  is the frequency of the temporal harmonic oscillations into which the motion of the particle can be imagined to be resolved. The corresponding transform relation is

$$S_{ij}(\xi) = \int_{-\infty}^{\infty} \Phi_{ij}(n) e^{i\xi n} dn, \quad \dots\dots\dots(2.9)$$

and in the particular case  $\xi = 0$

$$S_{ij}(0) = \overline{u_i u_j} = \int_{-\infty}^{\infty} \Phi_{ij}(n) dn,$$

showing that  $\Phi_{ij}(n)$  can be regarded as the density on the frequency axis of contribu-

tions to the stress tensor  $\overline{u_i u_j}$  (and in particular to the energy of the particle).

Substituting from (2.9) in (2.4) we find

$$\overline{X_i X_j} = \int_{-\infty}^{\infty} [\Phi_{ij}(n) + \Phi_{ji}(n)] \left( \frac{it}{n} - \frac{e^{int} - 1}{n^2} \right) dn \quad \dots\dots(2.10)$$

The relations (2.2) and (2.5) show that  $\Phi_{ji}(-n) = \Phi_{ij}^*(n) = \Phi_{ji}(-n)$  (i.e.  $\Phi_{ij}(n)$  is a Hermitian tensor) and so  $\Phi_{ij}(n) + \Phi_{ji}(n)$  is an even function of  $n$ . (2.10) therefore reduces to

$$\overline{X_i X_j} = \int_0^{\infty} [\Phi_{ij}(n) + \Phi_{ji}(n)] \frac{1 - \cos nt}{\frac{1}{2}n^2} dn \quad \dots\dots\dots(2.11)$$

Thus whereas the density on the positive frequency-axis of contributions to  $\overline{u_i u_j}$  (which is independent of  $t$ ) is  $\Phi_{ij}(n) + \Phi_{ji}(n)$ , the density of contributions to the dispersion  $\overline{X_i X_j}$  is equal to the same function multiplied by the time-dependent weighting factor  $(1 - \cos nt)/\frac{1}{2}n^2$ .

The special forms obtained for the dispersion when  $t$  is either large or small are now capable of interpretation in terms of the behaviour of the weighting factor. When  $t$  is small  $(1 - \cos nt)/\frac{1}{2}n^2$  approximates to  $t^2$  and is therefore independent of  $n$ , and oscillations of all frequencies contribute to the dispersion in the same way as they contribute to the stress. When  $t$  is large the slowest oscillations dominate the dispersion; from (2.11)

$$\begin{aligned} \overline{X_i X_j} &= 2t \int_0^{\infty} [\Phi_{ij}(n) + \Phi_{ji}(n)] \frac{1 - \cos nt}{(nt)^2} d(nt) \\ &\rightarrow 2t [\Phi_{ij}(0) + \Phi_{ji}(0)] \int_0^{\infty} \frac{1 - \cos \theta}{\theta^2} d\theta \text{ as } t \rightarrow \infty, \\ &= t\pi [\Phi_{ij}(0) + \Phi_{ji}(0)] \quad \dots\dots\dots(2.12) \end{aligned}$$

which in view of (2.8) is identical with the main term of (2.7). Oscillations having frequencies which are multiples of  $1/t$  make no contribution to the dispersion because they return the particle to its starting point after time  $t$ .

The expression (2.11) describing the contribution to the dispersion from oscillations of different period shows that at no time is the dispersion dominated by the effect of oscillations of small period. Consequently there is no possibility of applying the equilibrium theory of high frequency oscillations(1), which has proved so successful in the study of the dynamics of turbulent motion, to the dispersion of a single fluid particle about a fixed point. It will be shown in a later paper that it is, however, possible to apply this theory in the Lagrangian analysis of diffusion.

### III. THE PROBABILITY DISTRIBUTION OF $X_i(t)$

Having established a formula for the dispersion of a single fluid particle about its original position, we now enquire into the complete probability distribution (it

will be more convenient to discuss the probability density function or p.d.f.) of the displacement components  $X_i(t)$ . If the form of this p.d.f. at different times can be established, the spatial distribution of the mean concentration of the property carried by the fluid particles is then known, and there is the possibility of obtaining a differential equation for the propagation (in time) of this mean concentration.

Little can be said about the p.d.f. of  $X_i(t)$  on purely theoretical grounds. We notice firstly, that when  $t$  is sufficiently small the p.d.f. of  $X_i(t)$  is identical with that of  $u_i(0)$  since

$$X_i(t) = \int_0^t u_i(t') dt' \approx t u_i(0).$$

$u_i(0)$  is the velocity of the fluid particle at the instant of its selection, and so the p.d.f. of  $X_i(t)$  is identical with the p.d.f. of the  $i$ -component of velocity at a definite point in space. Many measurements of this latter p.d.f. (more particularly of the second, third, and fourth moments) have been made in isotropic turbulence(7) and it has been found to be Gaussian within the limits of experimental error.

Secondly, when  $t$  is sufficiently large,  $u_i(t)$  loses all correlation with  $u_i(0)$  and so  $X_i(t)$  is the integral of a function, the values of which have finite correlation with each other for neighbouring parts of the range of integration only. This suggests the application of some theorem like the Central Limit Theorem(8), which asserts that the sum of a series of  $n$  terms has a p.d.f. which tends to normality as  $n$  tends to infinity provided the terms of the series have p.d.f.'s satisfying certain general conditions and are statistically independent of each other. An extension to an integral, with a continuous integrand, is needed but it is not self-evident. Theorems which partially fill this need by generalizing to the case of a series in which the terms may be statistically connected have been given (for instance, by Bernstein 9), but the conditions of validity of these theorems are rather specialized and we have not sufficient information about the turbulent velocity  $u_i(t)$  to be able to say whether any of these modified Central Limit Theorems is applicable. At least we can say that probability theory suggests that the p.d.f. of  $X_i(t)$  approaches the Gaussian form as  $t$  tends to infinity.

Fortunately there is available a considerable amount of experimental data about the p.d.f. of  $X_i(t)$  and, as a result of some recent work, the information covers the complete range of  $t$ . In almost every case the p.d.f. has been found to be Gaussian within the limits of experimental error. Considering first "small" values of  $t$  (i.e. values of  $t$  for which  $S_{ij}(\xi) \approx \overline{u_i u_j}$ ), a direct determination of the p.d.f. of  $X_i(t)$  was made by Schubauer(10) by observing the lateral distribution of mean temperature behind a small heated wire stretched across a uniform stream on which an isotropic turbulent motion was superimposed. Since the line source of heat is effectively moving through the field of turbulence, this is a case of one-dimensional diffusion (in a direction normal to the cylinder axis and to the stream velocity) and the time variable is effectively replaced by the distance down-stream from the cylinder. Schubauer's observations were confined to distances down-stream for which the law (2.5) was valid (showing that  $t$  was "small"), and he found the temperature profiles across the heated wake to be accurately Gaussian in shape.



The same experimental arrangement<sup>†</sup> has recently been used by Collis(11), but with a smaller scale of turbulence to permit a greater range of values of  $t$  to be explored. Collis's observations of the lateral distribution of mean temperature were made at positions for which (2.5) was valid ("small" values of  $t$ ), at positions for which (2.7) was valid ("large" values of  $t$ ), and at intermediate positions. No significant deviation from a Gaussian temperature distribution was observed at any of these positions. As in Schubauer's experiments, the field of turbulence was generated by blowing air through a square mesh grid of bars.

Several other measurements over a limited range of values of  $t$  have been reported; for instance Kalinske and Pien(12) found a Gaussian distribution of  $X_i(t)$  at "large" times for which (2.7) was valid. These authors injected a mixture of hydrochloric acid and water into a stream of water in an open channel and observed the lateral distribution of average concentration down-stream from the source. The turbulence in the stream was fully developed but was approximately homogeneous near the centre of the cross-section.

The inference from all these experiments is that in all cases of homogeneous turbulence the probability distribution of  $X_i(t)$  is normal. Except at very large values of  $t$  this is probably an *approximate* result, although the approximation may be very close. All the experiments described were concerned with one-dimensional diffusion but the inference will apply to each of the components  $X_1, X_2, X_3$  independently, even though their mean squares have different values. It is an interesting speculation (of which no use will be made in this paper) from this result that the *joint* probability distribution of the velocities of the fluid particle at different times is Gaussian.  $X_i(t)$  can be written approximately as

$$\sum_{\tau=0}^N u_i(\tau\delta t)\delta t,$$

where  $(N+1)\delta t = t$ , and since the sum of terms which have a joint normal distribution is itself distributed normally, it follows from this speculation that  $X_i(t)$  is distributed normally. However, the converse cannot be implied and further evidence for the normality of the joint p.d. of the velocities  $u(t_1), u(t_2), \dots, u(t_n)$  at any  $n$  times must be sought before it can be assumed confidently. There are no known experimental results which conflict with this speculation.

#### IV. THE DIFFERENTIAL EQUATION FOR MEAN CONCENTRATION

As mentioned earlier, the probability distribution of  $X_i(t)$  is intimately connected with the spatial distribution of mean concentration of the property transported by the fluid particle. Let us imagine that at time  $t = 0$  the fluid within a certain region of space of volume  $V$  possesses with unit intensity the property which is being diffused; this defines the marked fluid. Then the problem is to find the distribution of average intensity or concentration (over a large number of trials) of this property at subsequent times. This is equivalent to finding the probability  $P(\mathbf{x}, t)$  that the point defined by the position vector  $\mathbf{x}$  lies within the marked fluid at time  $t$ , given that  $P(\mathbf{x}, 0)$  is unity or zero according as the point  $\mathbf{x}$  lies within or outside a given region of space of volume  $V$ . The volume of marked fluid is the same at all times



and so

$$\int P(\mathbf{x}, t) dv(\mathbf{x}) = V, \quad \dots\dots\dots(4.1)$$

where  $dv(\mathbf{x})$  is an element of volume at position  $\mathbf{x}$  and the integration is over all space.

Now each particle diffuses independently of its neighbours in the sense that the p.d.f. of its displacement is independent of the presence of neighbouring particles of marked fluid. Representing by  $Q(\mathbf{X}, t)$  the p.d.f. of the vector displacement  $\mathbf{X}$  of any fluid particle during a time  $t$ , the probability that a fluid particle at time  $t$  will lie within a volume element  $dv(\mathbf{X})$  whose centre is at position  $\mathbf{X}$  relative to the initial position of the particle is  $Q(\mathbf{X}, t) dv(\mathbf{X})$ . Hence the mean concentration for a finite volume of marked fluid is

$$P(\mathbf{x}, t) = \int P(\mathbf{x}', 0) Q(\mathbf{x} - \mathbf{x}', t) dv(\mathbf{x}') \quad \dots\dots\dots(4.2)$$

Special parameters describing the distribution of mean concentration are readily obtained from (4.2); for instance, the "centre" of the distribution of mean concentration at time  $t$  has coordinates

$$\begin{aligned} c_i(t) &= \frac{1}{V} \int x_i P(\mathbf{x}, t) dv(\mathbf{x}), \\ &= c_i(0) + \int (x_i - x'_i) Q(\mathbf{x} - \mathbf{x}', t) dv(\mathbf{x}'), \\ &= c_i(0) + \overline{X_i(t)} \quad \dots\dots\dots(4.3) \end{aligned}$$

Likewise the generalized dispersion of the mean concentration of marked fluid at time  $t$  is

$$\begin{aligned} d_{ij}(t) &= \frac{1}{V} \int (x_i - c_i)(x_j - c_j) P(\mathbf{x}, t) dv(\mathbf{x}) \\ &= d_{ij}(0) + \overline{X_i(t)X_j(t)} \quad \dots\dots\dots(4.4) \end{aligned}$$

Thus the changes in the position of the centre and in the dispersion of the mean concentration are independent of the shape of the cloud of marked fluid, and are identical with the corresponding quantities for a single fluid particle.

The experimental result of the previous section states that separate probability distributions of  $X_1$ ,  $X_2$ , and  $X_3$  are normal for all values of  $t$ . This is a very significant result because of the consequent simplicity in the way in which the average concentration of the diffusing property is propagated with time. It will be recalled that diffusion from an instantaneous point source which is described by the molecular diffusion equation also leads to concentration profiles which are Gaussian at each instant. We should not expect the ordinary diffusion equation with constant diffusion coefficient to apply without any modification whatever, since in the case of turbulent diffusion the time-dependence of the dispersion is not represented by (2.7) at all times, but we are prompted to look more closely into the use of the ordinary diffusion equation.

As they stand the experimental results do not provide sufficient information for us to write down an expression for  $Q(\mathbf{x}, t)$  because they say nothing about the probability distribution of, say,  $X_1$  for *given* values of  $X_2$  and  $X_3$ . In most cases of homogeneous turbulence the probability distributions of  $X_1$ ,  $X_2$ , and  $X_3$  will be approximately independent, in which case no further information is needed. However, in order to make the analysis as general as possible (consistent with obtaining a definite expression for  $Q$ ) we shall make the assumption\* that the distributions of  $X_1$ ,  $X_2$ , and  $X_3$  are jointly as well as separately normal. In this case it is known that the appropriate expression for  $Q$  is

$$Q(\mathbf{X}, t) = \frac{1}{(8\pi^3\Omega)^{\frac{1}{2}}} \exp \left[ -\frac{1}{2\Omega} w_{rs} X_r X_s \right], \quad \dots\dots\dots(4.5)$$

where

$$\Omega = \begin{vmatrix} \overline{X_1^2} & \overline{X_2 X_1} & \overline{X_3 X_1} \\ \overline{X_1 X_2} & \overline{X_2^2} & \overline{X_3 X_2} \\ \overline{X_1 X_3} & \overline{X_2 X_3} & \overline{X_3^2} \end{vmatrix}.$$

$w_{rs}$  is the co-factor of the typical element  $\Omega_{rs}$  of this determinant, and repeated suffixes are regarded as being summed over the values 1, 2, and 3. Hence (4.2) becomes

$$P(\mathbf{x}, t) = \int \frac{P(\mathbf{x}', 0)}{(8\pi^3\Omega)^{\frac{1}{2}}} \exp \left[ -\frac{1}{2\Omega} w_{rs}(x_r - x'_r)(x_s - x'_s) \right] d\mathbf{x}' \quad \dots\dots\dots(4.6)$$

The differential equation which we are encouraged to test as an alternative description of the diffusion is

$$\frac{\partial P(\mathbf{x}, t)}{\partial t} = K_{ij} \frac{\partial^2 P(\mathbf{x}, t)}{\partial x_i \partial x_j} \quad \dots\dots\dots(4.7)$$

It is, in fact, easily verified by substitution that (4.6) is a solution of the equation (4.7) provided the diffusion coefficient tensor  $K_{ij}$  is such as to satisfy the equations

$$K_{ij} w_{ij} = \frac{1}{2} \frac{d\Omega}{dt}, \quad \dots\dots\dots(4.8)$$

$$K_{ij} w_{ir} w_{js} = -\frac{1}{2} \Omega^2 \frac{d(w_{rs}/\Omega)}{dt} \quad \dots\dots\dots(4.9)$$

Remembering that by definition

$$\Omega_{ir} w_{jr} = \Omega \delta_{ij},$$

where  $\delta_{ij} = 1$  if  $i = j$ , and  $= 0$ ; otherwise, (4.9) can be reduced to

$$K_{ij} = \frac{1}{2} \frac{d\Omega_{ij}}{dt}, \quad \dots\dots\dots(4.10)$$

which is also consistent with (4.8). Hence the general coefficient of diffusion in

\*This assumption is not very restrictive. It would be an exact inference if measurements showed that the component of  $\mathbf{X}(t)$  in an *arbitrary* direction is normally distributed.

homogeneous turbulence is

$$K_{ij} = \frac{1}{2} \frac{d\overline{X_i X_j}}{dt} = \frac{1}{2} \int_0^t [S_{ij}(\xi) + S_{ji}(\xi)] d\xi \quad \dots\dots\dots (4.11)$$

Equivalently, the rate of transfer of the diffusing property across unit area normal to the  $x_i$  axis at time  $t$  is

$$K_{ij} \frac{\partial P(\mathbf{x}, t)}{\partial x_j}.$$

The generality of (4.7) and (4.11) will not often be needed, of course; if principal axes for the displacement tensor  $\overline{X_i X_j}$  are used,  $K_{ij} = 0$  when  $i \neq j$ , and if the turbulence has rotational symmetry,  $K_{11} = K_{22} = K_{33}$ , in which case (4.7) reduces to the ordinary molecular diffusion type of equation.

Thus the diffusion equation (4.7) (sometimes called a Fickian equation)\* provides a description of the diffusion process when Eulerian analysis is used, provided the coefficient  $K_{ij}$  varies with time according to (4.11).  $K_{ij}$  is initially zero, increases with time at first linearly and then more slowly, and finally tends to a constant value. It has long been appreciated (see Dryden 13) that an equation of type (4.7) would represent the diffusion process when  $t$  is "large," since the time-dependence of the dispersion (as in (2.7)) then has the same form as in the case of molecular diffusion, but it has now been shown that (4.7) is valid at all times. The special character of the time  $t = 0$ , when the diffusion coefficient is zero, lies in the fact that the position of the marked fluid is then known with certainty, i.e.  $P(\mathbf{x}, 0)$  has either of the values 0 or 1. The diffusion gradually changes character as the velocity of a marked particle loses correlation with its value at  $t = 0$ , and the random variations of the velocity of a particle increasingly dominate the diffusion.

It is worth while to notice that the validity of (4.7) for all values of  $t$  is, in a sense, an accident. It so happens that  $X_i(t)$  has a Gaussian probability distribution at all values of  $t$ , for reasons which vary from one range of  $t$  to the next. In molecular diffusion, the appropriate differential equation (if there is one) at values of  $t$  which are not "large" is of no particular interest in view of the small time scale of molecular motion under normal conditions. However, with turbulent diffusion, values of  $t$  which are not "large" will very often be those with which we are most concerned, and it is a fortunate accident that a single differential equation describes the complete process with high accuracy. In other words, we should treat (4.7) as a phenomenological and not as a fundamental law of turbulent diffusion.

The fortunate circumstance that  $X_i(t)$  has a Gaussian p.d.f. at all times ceases to apply accurately when the turbulence is not homogeneous, since the p.d.f. of  $u_i(o)$  is influenced by the spatial variation of the turbulence. This point is illustrated well in some measurements made by Skramstad, and recorded by Dryden(13), of the temperature distribution down-stream from a heated wire placed in the boundary layer close to a flat surface. The width (in a direction perpendicular to the surface) of the heat wake at four positions down-stream showed a linear in-

\*The author has been unable to find the derivation of this term.



crease with distance, so that the common distribution of temperature at these four positions was identical with the p.d.f. of the component of turbulent velocity at right angles to the surface. Dryden's diagram shows the distribution to be clearly not Gaussian; the diffusion is faster in the direction away from the boundary as is to be expected from the greater intensity of turbulence near the boundary. Thus in such a case the equation (4.7) is not valid; nor is it necessary that any other differential equation should describe the spread of heat when  $t$  is "small."

The conclusions established in this section are similar to assumptions about turbulent diffusion in the atmosphere which have been made by Sutton(4). Sutton assumes that the distribution of mean concentration is Gaussian at all values of  $t$  and adopts a form of (one component of) the function  $S_{ij}(\xi)$  which is such that

$$\int_0^t S_{ij}(\xi) d\xi$$

increases indefinitely with  $t$ . This has the effect of postponing indefinitely the asymptotic condition (2.7). Equation (2.12) shows that the asymptotic condition is reached only when  $\Phi_{ij}(1/t) \approx \Phi_{ij}(0)$ , so that in the atmosphere where the spectrum function  $\Phi_{ij}(n)$  varies appreciably for small values of  $n$  (corresponding to a wide range of slow eddies), the required value of  $t$  is indeed likely to be very large. Sutton introduces this form of the function  $S_{ij}(\xi)$  on the grounds that the size of the effective eddy governing the diffusion is determined by the dimensions of the cluster of marked particles and this latter quantity increases with time. This seems to be a confusion of the two types of problem, diffusion from a fixed centre and diffusion of a moving cloud, or, as they have been termed here, Eulerian and Lagrangian types. In an Eulerian analysis, which in effect Sutton uses, the size of the cloud is not relevant; as shown above the whole analysis is in fact based on the behaviour of a single particle. The increase of  $K_{ij}$  with time is due, not to an increase in the mean width of a cloud which may happen to be present, but to the fact that velocity oscillations of low frequency are becoming more and more effective in dispersing each particle about its original position. Allowance for the size of the cloud is certainly necessary in a Lagrangian analysis, as has been pointed out by Richardson (3) and as will be discussed in the later paper.

## V. USE OF THE DIFFUSION EQUATION

In principle it is now possible to solve any problem of the diffusion of a given initial distribution of marked fluid in a field of homogeneous turbulence. Of course in practice there will usually be the difficulty that the correlation tensor  $S_{ij}(\xi)$  is an unknown function of  $\xi$ , but this function affects only the time-dependence of the turbulent diffusion coefficient, i.e. it does not affect the spatial distribution of concentration but determines the instants at which the various spatial distributions occur. The most serious limitation to the applicability of equation (4.7) is the requirement that the turbulence should be homogeneous, since in most practical problems the lack of spatial uniformity is very marked. Cases which seem to offer scope for application are wind tunnel streams, the oceans, and the atmosphere



(excluding the layer near the ground) but in none of these cases have many suitable observations of diffusion been made.

If equation (4.7) is to be used, it is necessary to know with certainty, and not as an average, the initial spatial distribution of the diffusing property. Equation (4.7) is linear, so that the diffusion of a property like temperature which varies continuously in space can be handled by regarding the initial distribution of temperature as composed of a number of isothermal shells of small thickness, each of which diffuses independently. Thus we can replace the function  $P(\mathbf{x}, t)$  by the more general function  $I(\mathbf{x}, t)$  representing average intensity of the property which is being diffused; at  $t = 0$ ,  $I$  is a known non-fluctuating function of  $\mathbf{x}$ .

In the case of problems in which there is a continuously emitting source and the distribution of intensity is independent of time, the appropriate value of the diffusion coefficient will depend on position. This will also be true in the case of (idealized) steady distribution problems in which the fluid instantaneously occupying certain regions of space is held at a given intensity. Both of these classes of problem should be handled as resultants of the superposition of a number of instantaneous source problems, to each of which the solution (4.6) is applicable. We note that the transfer of intensity at parts of the field which are far distant from a source will be contributed chiefly by fluid particles which have lost all correlation with their motion when they were close to the source, and hence in these distant regions the transfer coefficient will be approximately independent of time and position. The governing equation will then be

$$K_{ij} \frac{\partial^2 I(\mathbf{x})}{\partial x_i \partial x_j} = 0 \quad \dots\dots\dots (5.1)$$

If fluid particles take a time of order  $T$  to lose correlations with their initial motion, i.e.  $S_{ij}(T) \approx 0$ , a particle travels a distance of order  $T\sqrt{u_i^2}$  in this time and so the equation (5.1) will be valid in steady distribution problems at positions whose distance from the nearest source is at least of order  $T\sqrt{u_i^2}$ .

An experiment to which (5.1) is roughly applicable has been carried out by Taylor(14) who measured the temperature distribution in the annulus between two heated concentric cylinders, the inner of which rotated at a speed far above that required to produce turbulent flow. The distribution of mean circumferential velocity is not uniform here, but we are concerned only with radial diffusion. Measurements of the turbulent velocities between concentric cylinders at speeds above and below that used by Taylor have been made by MacPhail(15), and while these show that the mean square of the radial velocity is certainly not constant across the annulus, it has a stationary value near the centre and varies by about 30 per cent. over the central 50 per cent. of the annulus. It seems probable that the scale of the turbulence is of the order of a fraction (say 1/10) of the annulus width and this we may estimate to be also the radial distance travelled by a fluid particle before losing nearly all correlation with its initial motion. Thus in the central 50 per cent. of the annulus the radial turbulent motion is approximately homogeneous and the diffusion coefficient is approximately independent of time and position. Equation

(5.1) is then approximately valid and, transforming to cylindrical polar coordinates, the first integral is

$$r \frac{dI(r)}{dr} = \text{constant} \quad \dots\dots\dots(5.2)$$

This variation of intensity is identical with the distribution of temperature found by Taylor to hold over slightly more than the central 50 per cent. of the annulus: the agreement between (5.2) and the measurements is in fact better than would be expected from the approximations involved in the application of the theory.

## VI. ACKNOWLEDGMENT

This paper is published with the permission of the Chief Scientist, Ministry of Supply, England.

## VII. REFERENCES

- (1) BATCHELOR, G. K.—*Proc. 7th Int. Cong. Appl. Mech. London (1948), Introductory Volume.*
- (2) TAYLOR, G. I.—*Proc. Lond. Math. Soc.* **20** : 196 (1921).
- (3) RICHARDSON, L. F.—*Proc. Roy. Soc. A* **110** : 709 (1926).
- (4) SUTTON, O. G.—*Ibid.* **135** : 143 (1932).
- (5) KAMPÉ DE FÉRIET, J.—*Ann. Soc. Sci. Bruxelles* **59** : 145 (1939).
- (6) BATCHELOR, G. K.—*Proc. Roy. Soc. A* **195** : 513 (1949).
- (7) TOWNSEND, A. A.—*Proc. Camb. Phil. Soc.* **43** : 560 (1947).
- (8) CRAMER, H.—“*Mathematical Methods of Statistics.*” (Univ. Press: Princeton, 1946.)
- (9) BERNSTEIN, S.—*Math. Ann.* **47** : 1 (1926).
- (10) SCHUBAUER, G. B.—N.A.C.A. Tech. Rep. No. 524 (1935).
- (11) COLLIS, D. C.—Coun. Sci. Industr. Res. Aust. Div. Aero. Rep. No. A55 (1948).
- (12) KALINSKE, A. A., and PIEN, C. L.—*Industr. Engng. Chem.* **36** : 220 (1944).
- (13) DRYDEN, H. —*Ibid.* **31** : 416 (1939).
- (14) TAYLOR, G. I.—*Proc. Roy. Soc. A* **151** : 494 (1935).
- (15) MACPHAIL, D. C.—*Proc. 6th Int. Cong. Appl. Mech. Paris (1946).*

# THE FULLY DEVELOPED TURBULENT WAKE OF A CIRCULAR CYLINDER

By A. A. TOWNSEND\*

[*Manuscript received May 12, 1949*]

## *Summary*

Extending previous work on turbulent diffusion in the wake of a circular cylinder, a series of measurements have been made of the turbulent transport of mean stream momentum, turbulent energy, and heat in the wake of a cylinder of 0.159 cm. diameter, placed in an air-stream of velocity 1280 cm. sec.<sup>-1</sup>. It has been possible to extend the measurements to 950 diameters down-stream from the cylinder, and it is found that, at distances in excess of 500 diameters, the requirements of dynamical similarity are very nearly satisfied. To account for the observed rates of transport of turbulent energy and heat, it is necessary that only part of this transport be due to bulk convection by the slow large-scale motion of the jets of turbulent fluid emitted by the central, fully turbulent core of the wake, which had been supposed previously to perform most of the transport. The remainder of the transport is carried out by the small-scale diffusive motion of the turbulent eddies within the jets, and may be described by assigning diffusion coefficients to the turbulent fluid. It is found that the diffusion coefficients for momentum and heat are approximately equal, but that for turbulent energy is considerably smaller. On the basis of these hypotheses, it is possible to calculate the form of the mean velocity distribution in good agreement with experiment, and to give a qualitative explanation of the apparently more rapid diffusion of heat.

## I. INTRODUCTION

Since the original formulation of the mixing-length theory of turbulent diffusion by Prandtl and G. I. Taylor, this theory has been applied very extensively to the calculation of the mean properties of turbulent shear flows, and there is no doubt that it has provided a very useful framework for the interpretation and analysis of experimental results. Although the theory in its various forms has been moderately successful in its particular field of the prediction of overall results of turbulent diffusion, it is generally realized that the underlying physical picture of the diffusion process is untenable, and lately a number of attempts have been made to apply the methods of the statistical theory of turbulence to the problem of turbulent shear flow. If this problem is approached directly from the Navier-Stokes equations of motion, it is obviously necessary to make a considerable number of simplifying assumptions before any progress is possible. Chou (1, 2) and his co-workers (3) have chosen this course, and, by making a number of assumptions, some with experimental foundation, they have succeeded in obtaining solutions for a number of flows. In the particular example of wake flow, the predicted distribution of turbulent intensity is very different from that observed experimentally, and the theory

\*Emmanuel College, Cambridge.



needs modification before it may be considered a satisfactory description of conditions in the wake. It appears more profitable to seek the underlying regularities in the turbulent flow in the first place by experimental study of the statistical properties of the motion, and to use the regularities, when they are discovered, as the basis of a general theory. To this end, systematic measurements have been made in the circular jet by Corrsin (4, 5), in the two-dimensional channel and in the half-jet by Liepmann and Laufer (6, 7), in the turbulent boundary layer by Dryden (8), and in the cylinder wake by the author (9, 10). In this work, the author investigated the occurrence of intermittently turbulent flow in the wake of a circular cylinder by measuring triple and quadruple velocity correlations at a point, and, from these results, inferred the existence of a method of diffusion, whose magnitude is independent of the local intensity gradient, and which cannot be described by a diffusion coefficient. Briefly, the central, fully turbulent core of the wake is found to emit slow billows or jets of fully turbulent fluid, which, as a consequence of their motion away from the wake centre, transport turbulent energy, heat, or other transferable quantities at a rate depending on the local intensity and the outward velocity of the jet. Inside each jet, the turbulent motion is in the state of equilibrium postulated by Kolmogoroff in his theory of local isotropy, and consequently the distribution and properties of the small eddies are dependent only on the local energy dissipation and on the kinematic viscosity. These observations were made at positions from 80 to 160 diameters down-stream from the cylinder, where it is known that the character of the wake motion is still changing, so, to confirm the existence and importance of this mode of turbulent transport, an extended series of measurements has been completed at distances of 500 to 950 diameters from the cylinder. It became clear in the course of this work that, while the wake structure is not qualitatively different in these conditions, transport by jet motion is not solely responsible for the observed turbulent transfer, and ordinary diffusion down an intensity gradient may also be appreciable. The final description of the wake motion is still incomplete, but it appears to be self-consistent, and in time it may be possible to construct an exact theory of the wake based on the described structure.

## II. NOTATION

Consider the wake of a circular cylinder of diameter,  $d$ , placed in an air-stream of velocity,  $U_0$ . Then establish a coordinate system with origin on the axis of the cylinder,  $Ox$  parallel to the direction of  $U_0$ ,  $Oz$  along the axis of the cylinder, and  $Oy$  at right angles to  $U_0$  and to the cylinder axis. In the usual notation, the components of mean velocity at any point are  $U$ ,  $V$ ,  $W$ , and the instantaneous velocity components are  $U + u$ ,  $V + v$ ,  $W + w$ . The instantaneous pressure fluctuation from the mean value  $P$  is denoted by  $p$ , and the instantaneous temperature fluctuation from the mean value  $T$  (degrees absolute) by  $\theta$ .

Air density is denoted by  $\rho$ , kinematic viscosity by  $\nu$ , and thermometric conductivity by  $\kappa$ .

All mean values are taken with respect to time at a point fixed with respect to the cylinder, unless they are specifically defined differently.



## III. EXPERIMENTAL ARRANGEMENTS

The small wind-tunnel in the Cavendish Laboratory, Cambridge, was used in these experiments. The working section is of square cross-section, 15 by 15 in., and the free-tunnel turbulence levels are

$$\sqrt{u^2}/U_0 = 0.0006, \quad \sqrt{v^2}/U_0 = \sqrt{w^2}/U_0 = 0.0015.$$

The main series of measurements were made in the wake of a constantan wire, 0.159 cm. diameter, stretched centrally across the entrance to the working section. For measurements of heat diffusion, it was heated by passing through it a current of approximately 30 amperes, thus raising the air temperature in its neighbourhood by roughly 10°C. All the measurements were made at a wind-speed of 1280 cm.sec.<sup>-1</sup>, and for the 0.159 cm. diameter cylinder, the Reynolds number,  $(U_0 d)/\nu$  is 1360. A few measurements were made using a cylinder of 0.953 cm. diameter, at a Reynolds number of 8200.

Mean velocity in the wake was measured with a small total head tube of 1 mm. outside diameter. The mean temperature was measured with a copper-constantan thermocouple. The thermojunction wires were of 46 S.W.G., and the cold junction was placed in the air-stream, alongside the hot junction, but clear of the wake. The observed rises of air temperature were of the order 0.3°C.

The hot-wire anemometer, in both the single- and double-wire forms, was used for all the turbulence measurements. By using an X-form wire in a bridge circuit with three branches, it is possible to obtain simultaneously separate electrical signals proportional to the down-stream and cross-stream components of the turbulent motion. Then, using the method previously described by the author (11), it is possible to measure the triple correlation coefficients  $\overline{u^2 v}/\overline{u^2} \sqrt{\overline{v^2}}$  and  $\overline{v^3}/(\overline{v^2})^{3/2}$ , and, from knowledge of the turbulent intensities, to calculate the diffusion rates  $\overline{u^2 v}$  and  $\overline{v^3}$ . To measure the remaining diffusion rate  $\overline{vw^2}$ , an X-form wire is set up so that the plane of the wires is at 45° to the axis of the cylinder and parallel to the  $Ox$  axis. Then the output is proportional to  $(v + w)$ , and the correlation  $\overline{(v + w)^3}/[\overline{(v + w)^2}]^{3/2}$  may be measured.

Since

$$\overline{(v + w)^2} = \overline{v^2} + \overline{w^2}$$

and

$$\overline{(v + w)^3} = \overline{v^3} + 3\overline{vw^2}$$

in consequence of the symmetrical properties of the wake,  $\overline{vw^2}$  may be calculated from previous knowledge of  $\overline{v^3}$ .

The turbulent energy dissipation has been found by measuring  $(\overline{\partial u/\partial x})^2$  using the technique of time differentiation of the hot-wire signal, and then assuming the existence of local isotropy, which leads to a value of the mean energy dissipation of

$$\overline{W} = 15\rho\nu \overline{(\partial u/\partial x)^2}.$$

The intermittency factor,  $\gamma$ , defined as the probability that the flow at a given point shall be turbulent at any instant, has been calculated from the flattening factors of velocity derivatives as described in a previous paper (12).

The turbulent transport of heat has been measured directly following the method suggested by Corrsin (13), which depends on the fact that a hot-wire responds both to fluctuations in air velocity and to fluctuations in air temperature, but that the



## IV. DYNAMICAL SIMILARITY OF FLOW IN THE WAKE

It has often been assumed that, sufficiently far down-stream, the processes of a wake or other shear flow will settle down to a statistical equilibrium in which the velocity pattern is self-perpetuating, merely changing scale and intensity with

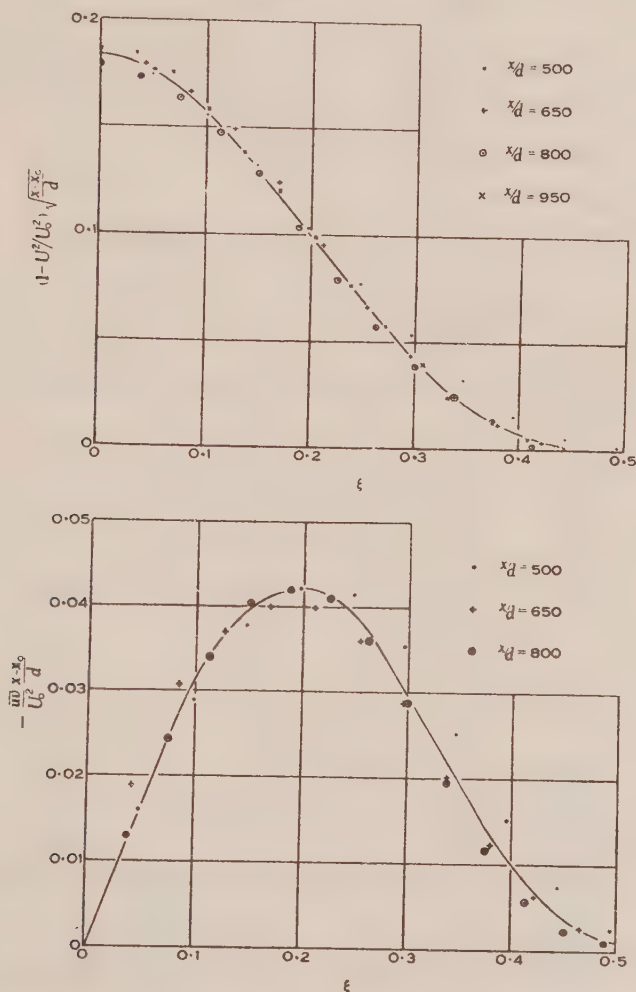


Fig. 2.—Mean velocity and shear stress.

distance from the cylinder. Then any mean property of the wake, expressed non-dimensionally in terms of the local characteristic length,  $l_1$ , and velocity,  $v_1$  (which are functions of  $x$  only), is a universal function of  $y/l_1$ . By applying dimensional analysis to the wake, this leads to the following relations :

$$l_1 \propto (x - x_0)^{1/2}$$

$$v_1 \propto (x - x_0)^{-1/2},$$

where  $x_0$  is the virtual origin of the wake, and then to

$$1 - U^2/U_0^2 = X^{-1/2}f_1(\xi)$$

$$T/T_a - 1 = X^{-1/2}g_1(\xi)$$

$$\overline{uv}/U_0^2 = -X^{-1}f_2(\xi)$$

$$\overline{v\theta}/U_0T_a = X^{-1}g_2(\xi)$$

$$\overline{u^2}/U_0^2 = X^{-1}h_1(\xi), \quad \overline{v^2}/U_0^2 = X^{-1}h_2(\xi), \quad \overline{w^2}/U_0^2 = X^{-1}h_3(\xi)$$

$$\overline{u^2v}/U_0^3 = X^{-3/2}j_1(\xi), \quad \overline{v^3}/U_0^3 = X^{-3/2}j_2(\xi), \quad \overline{vw^2}/U_0^3 = X^{-3/2}j_3(\xi),$$

where  $X = (x - x_0)/d$ ,  $Y = y/d$ , and  $\xi = YX^{-1/2}$ .

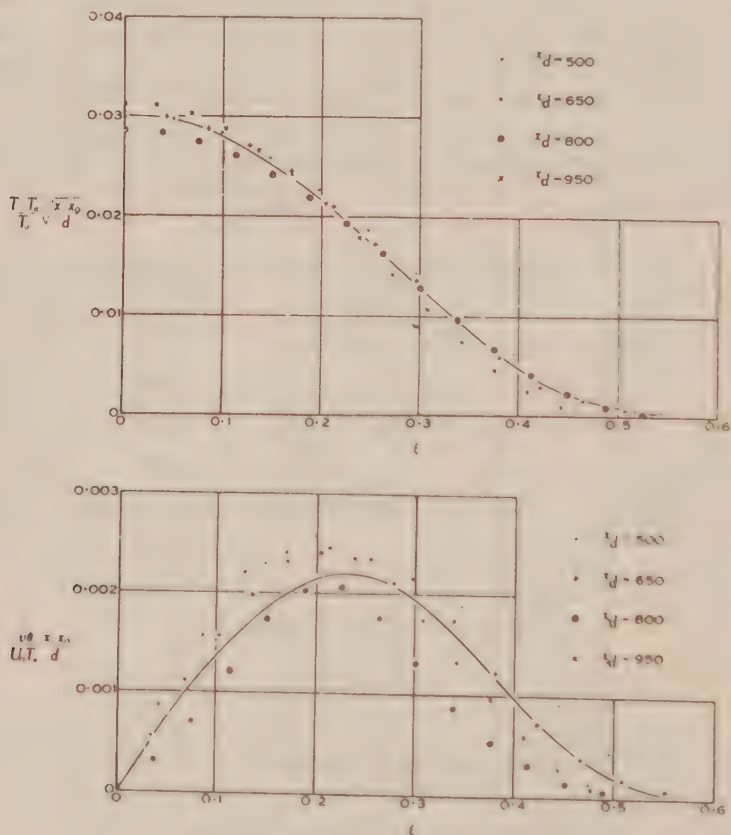


Fig. 3.—Mean temperature and heat flux.

$f_1$ ,  $f_2$ , etc. are universal functions at a particular cylinder Reynolds number, and heat input. Although the hypothesis of dynamical similarity is the basis of all existing theories of wake flow, it has been shown that it can only be true at distances in excess of 500 diameters down-stream from the cylinder (9). The present results in the range 500 to 950 diameters down-stream may fall in a range within which similarity is approximately valid, and observed values of the various quantities are



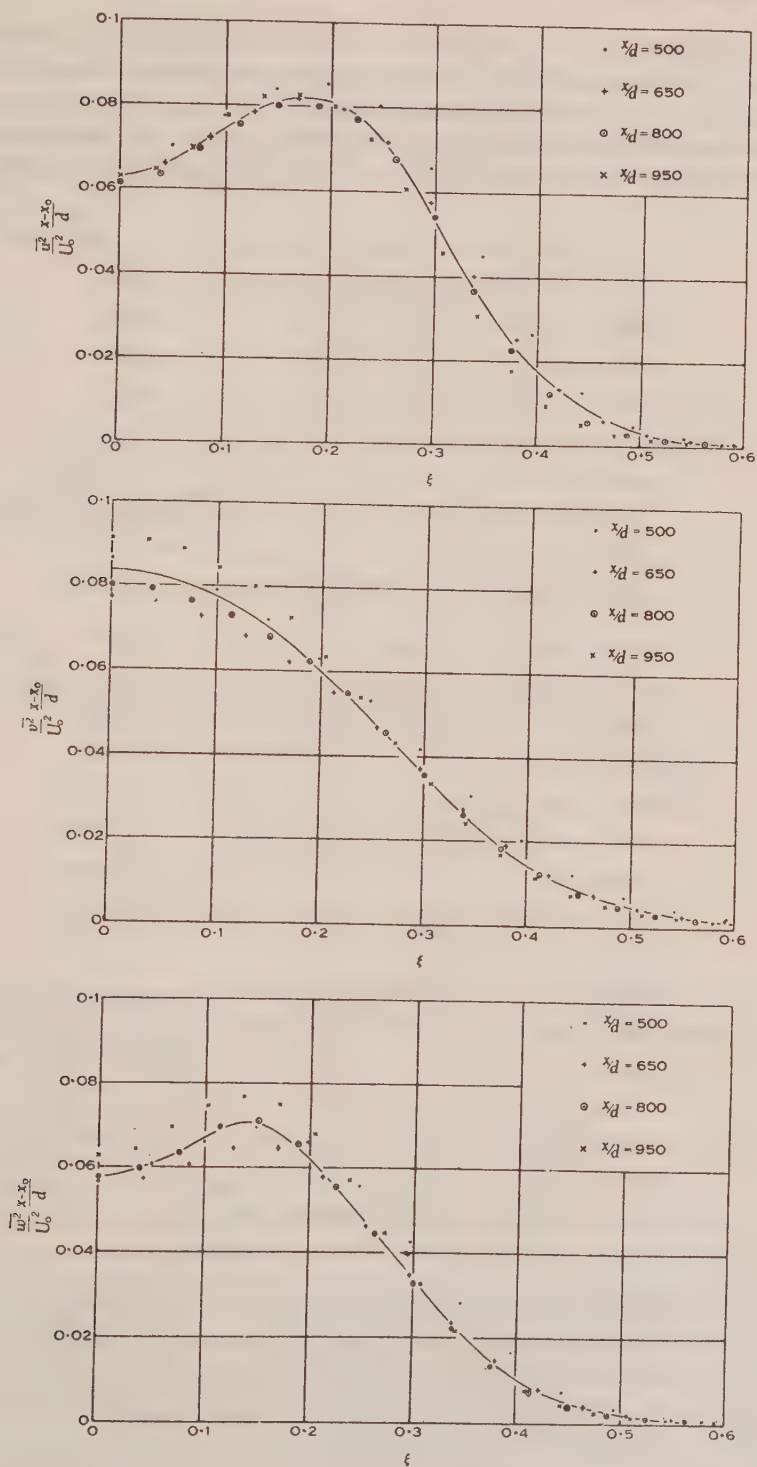


Fig. 4.—Intensities of turbulence components.

plotted in Figures 2-7 to find whether universal functions can be found to fit the observations. The only arbitrary constant is  $x_0/d$  which has been set at  $+90$ , and the diagrams show that, plotted in this way against  $\xi$ , the measured quantities from traverses at the four positions,  $x/d = 500, 650, 800$ , and  $950$ , fall close to the mean curves representing the universal functions. There are small but significant de-

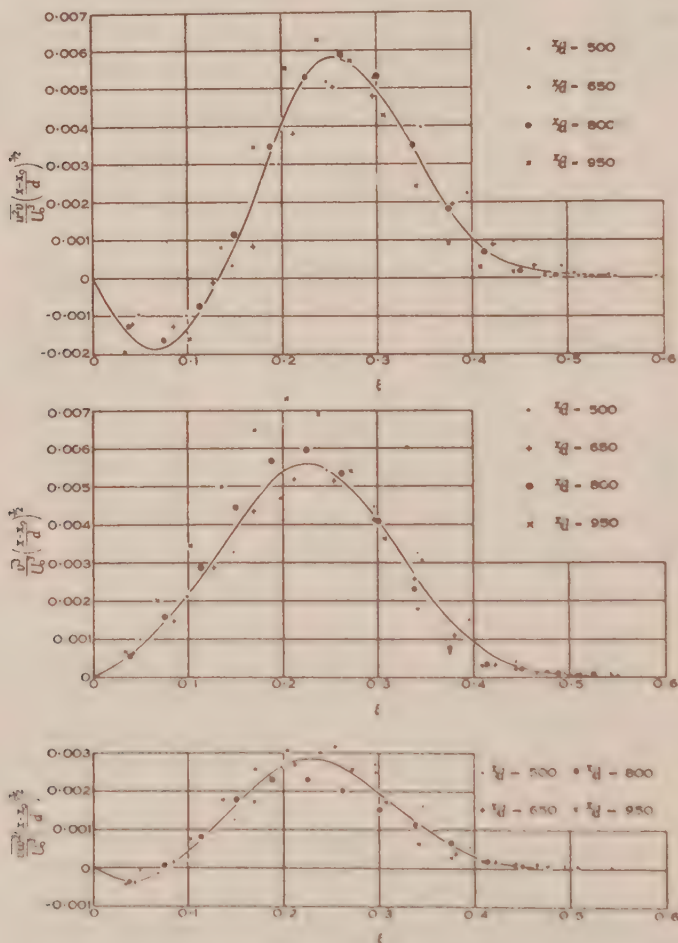


Fig. 5.—Energy diffusion rates.

viations from similarity at the extreme edge of the wake, in particular at  $x/d = 500$ , and these are presumably related to the initial excess of turbulent motion (10), and the consequent over-rapid diffusion in the early stages of wake development. The deviations over most of the wake are sufficiently small to justify the assumption of similarity in the analysis of the results.

When this assumption is made, the equations of motion for the wake can be expressed in terms of  $\xi$  only. Making the usual boundary layer assumptions, the

equation of mean motion is

$$U_0 \frac{\partial U}{\partial x} = - \frac{\partial \overline{uv}}{\partial y} + \nu \frac{\partial^2 U}{\partial y^2},$$

which becomes

$$\frac{1}{2} \xi f_1 = f_2 + \frac{1}{2} \frac{\nu}{U_0 d} f'_1 \quad \dots\dots\dots(1)$$

Similarly, the heat transport equation

$$U_0 \frac{\partial T}{\partial x} = - \frac{\partial \overline{v\theta}}{\partial y} + \kappa \frac{\partial^2 T}{\partial y^2},$$

becomes

$$\frac{1}{2} \xi g_1 = g_2 + \frac{\kappa}{U_0 d} g'_1 \quad \dots\dots\dots(2)$$

Finally, the energy equation

$$15\nu \left( \frac{\partial \overline{u}}{\partial x} \right)^2 + \frac{1}{2} U_0 \frac{\partial}{\partial x} (\overline{u^2} + \overline{v^2} + \overline{w^2}) + \overline{uv} \frac{\partial U}{\partial y} \\ + \frac{1}{2} \frac{\partial}{\partial y} (\overline{u^2 v} + \overline{v^3} + \overline{vw^2}) + \frac{1}{\rho} \frac{\partial}{\partial y} (\overline{pv}) = 0,$$

becomes

$$15 \frac{\nu}{U_0 d} \kappa - \frac{1}{2} [h_1 + h_2 + h_3 + \frac{1}{2} \xi (h'_1 + h'_2 + h'_3)] + \frac{1}{2} f_2 f'_1 \\ + \frac{1}{2} (j'_1 + j'_2 + j'_3) + l' = 0, \quad \dots\dots\dots(3)$$

where

$$\kappa(\xi) = X^2 \frac{d^2}{U_0^2} \left( \frac{\partial \overline{u}}{\partial x} \right)^2$$

and

$$l(\xi) = X^{3/2} \frac{\overline{pv}}{\rho U_0^3}$$

are two more universal functions, representing viscous dissipation and pressure flow. The various terms in this equation are plotted in Figure 8, the pressure term being obtained by difference.

## V. ENERGY AND HEAT TRANSPORT: JET CONVECTION AND GRADIENT DIFFUSION

Recently, in an investigation of the turbulent flow at distances between 80 and 160 diameters from a circular cylinder (10), the author concluded that most of the turbulent energy transport is due to slow, large-scale movements which convect turbulent energy at a rate not dependent on the local intensity gradient. The main reason for this conclusion was that, although the mean local intensity gradients are always small and positive over much of the wake, the transport rates are large and positive, i.e. transport occurs in considerable quantities against the intensity gradient. This anomalous transport was explained in terms of convection of turbulent fluid by the motion of billows or jets emerging from the fully turbulent central core of the wake, a motion always away from the wake centre. The present experiments

have been carried out at much greater distances down-stream from the cylinder, where it is found that, close to the wake centre, substantial energy transport inwards occurs (Fig. 5), and consequently diffusion down the intensity gradient must be taking place. Nevertheless, it is not possible to explain the results in terms of gradient diffusion alone, and a considerable part of the energy transport must be ascribed to a jet convection of the type described. This is most easily seen by

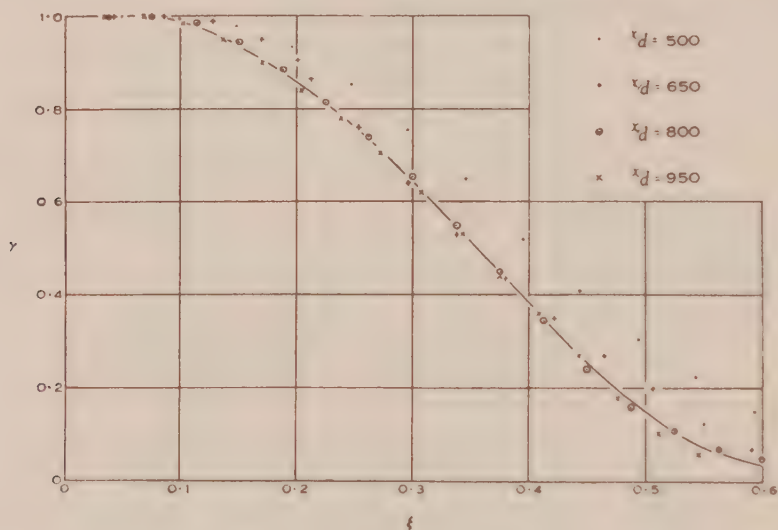


Fig. 6.—Intermittency factors.

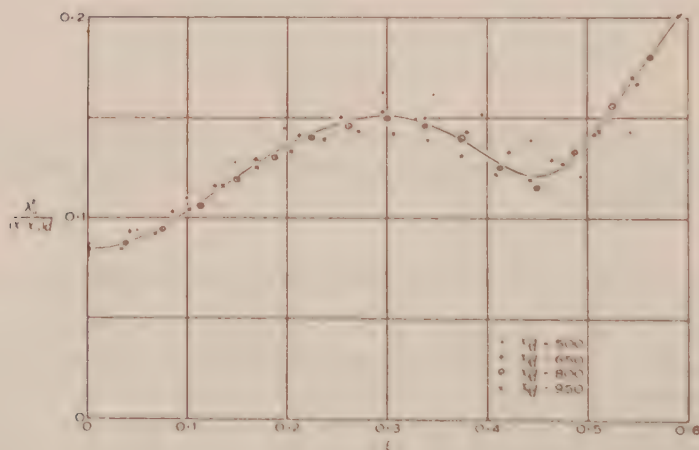


Fig. 7.—Dissipation length parameter.

comparing the gradient of  $(\overline{u^2} + \overline{v^2} + \overline{w^2})/\gamma$ , the mean turbulent intensity within the jets, with  $(\overline{u^2\nu} + \overline{v^3} + \overline{vw^2})/\gamma$ , the mean rate of energy transport within the jets. Over an appreciable part of the wake, the transport is still against the local intensity



gradient. Also, although heat transport is always outwards and down the local temperature gradient, in order to explain the transport as a simple gradient diffusion it is necessary to postulate anomalously high and variable eddy diffusion coefficients, and it is much more credible to suppose that both gradient diffusion and jet convection are occurring simultaneously.

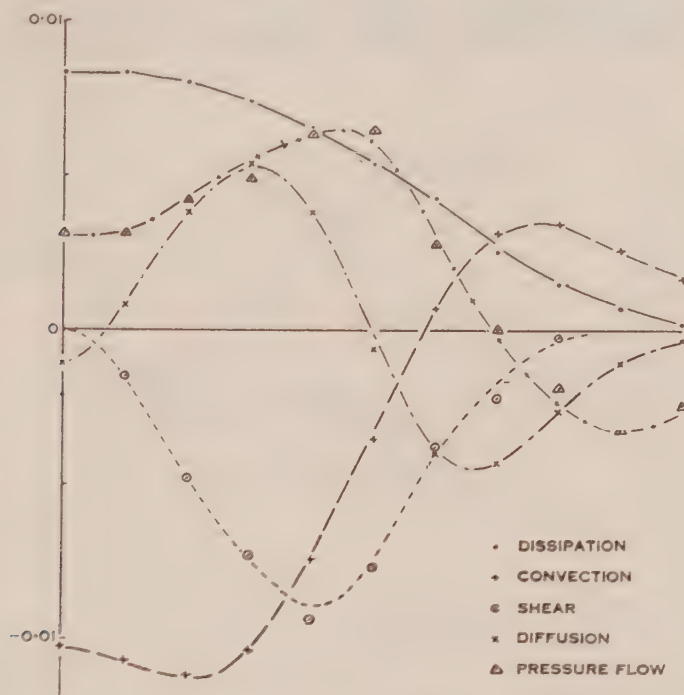


Fig. 8.—Energy balance.

Let us now consider the results and attempt to estimate the contributions of jet convection and gradient diffusion to the total energy and heat transport. It will be assumed that the turbulent motion in the wake consists of a large-scale motion, i.e. the jet motion, whose scale is comparable with the wake width, superimposed on a motion of scale small compared with the wake width, which contains the bulk of the turbulent energy. Then the effect of the jet motion is to convect turbulent energy or heat away from the wake centre, while the small-scale motion diffuses momentum, energy, and heat down the local intensity gradients. The difference in scale is supposed to be such that these processes occur independently of one another. Then a quantity which, like heat and turbulent intensity can only be transferred to turbulent fluid, takes a value different from its ambient value only within the regions of turbulent motion, and the total rate of turbulent transport may be written

$$\overline{vM} = \overline{VM} - \gamma \kappa_M \frac{\partial}{\partial y} \left( \frac{\overline{M}}{\gamma} \right),$$

where  $V$  is that part of  $v$  which is due to the large-scale jet motion only,  
 $M$  is the instantaneous value of the quantity,  
 $\gamma$  is the intermittency factor defined above,  
 $\kappa_M$  is the eddy diffusion coefficient for the quantity.

The first term represents the jet convection, and the second the gradient diffusion, expressed in terms of the mean intensity within the turbulent fluid,  $\bar{M}/\gamma$ . The

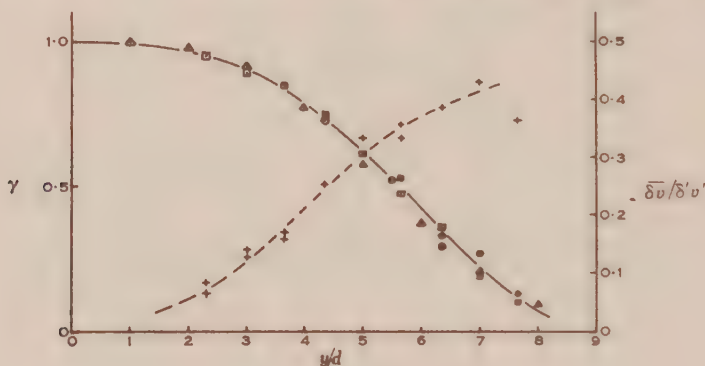


Fig. 9A.— $\delta$  measurements.

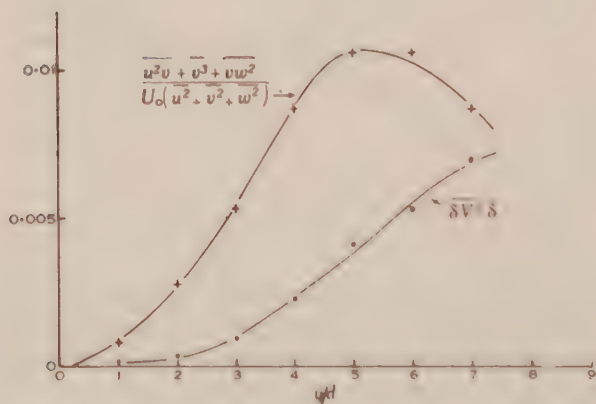


Fig. 9B.—Convection of energy.

direct approach is to attempt the measurement of the jet convection term,  $V\bar{M}$ , but this is not very easy. In an earlier paper, the author wrote the jet convection term as the product of the mean bulk velocity of the turbulent fluid, the mean intensity of the quantity within the jets and the intermittency factor

$$\overline{VM} = \bar{\delta V} / \bar{\delta} \times \bar{M} / \gamma \times \gamma.$$

It is assumed in this that no systematic variations of  $V$  or  $M$  take place across the jet, and it is now realized that this is unlikely to be true. Nevertheless, some meas-

measurements of  $\overline{\delta V}$  have been made using the  $\delta$  signal for a cylinder Reynolds number of 8200. At lower Reynolds numbers, the discrimination between turbulent and non-turbulent fluid is not sufficiently rapid and accurate, but at this Reynolds number, consistent results can be obtained both for  $\gamma$  and  $\overline{\delta V}$ . The measurements of  $\gamma$  and  $\overline{\delta V}$  are given in Figure 9, together with energy diffusion results. Only near the edge of the wake is the term  $\overline{\delta V M}/\gamma$  comparable with the total diffusion, and it seems that the original assumption is badly in error near the wake centre.

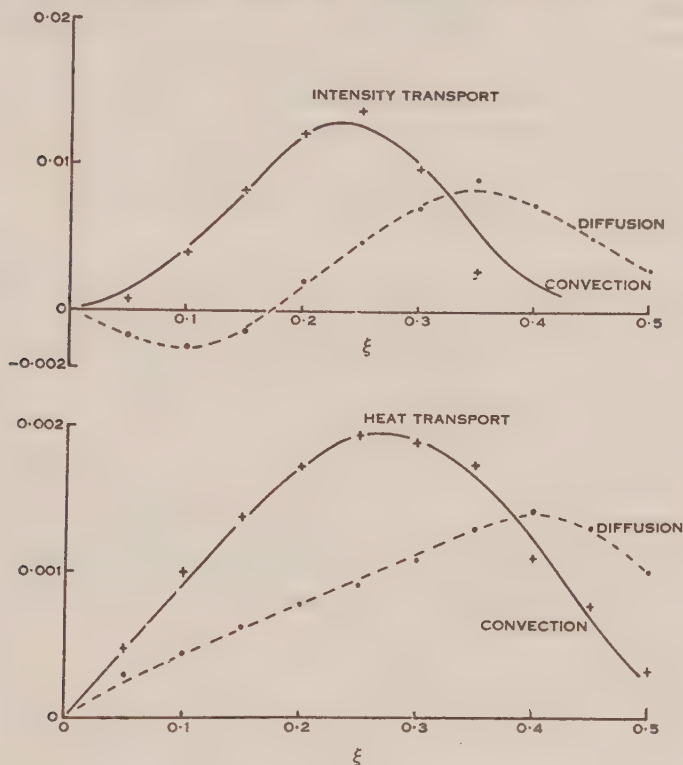


Fig. 10.—Jet convection and gradient diffusion.

It is easy to see the reason for this. The large-scale motion represented by  $V$  can have no large velocity gradients, and has a spatial mean value of zero. This implies that, where  $\gamma$  is large, turbulent fluid is flowing away from the wake centre only near the centre of an individual jet, and near the edges the flow must be toward the centre. If, as is very probable, the intensity of the diffusing quantity also reaches a maximum near the centre of the jet, then

$$\overline{VM} \gg \frac{\overline{\delta V}}{\delta} \overline{M}.$$

On the other hand, when  $\gamma$  is small, the variation of  $V$  across the jet will be small, and  $\overline{\delta V M}/\gamma$  is a good approximation to  $\overline{VM}$ . This is in fact observed, that, when

$\gamma$  is small, direct measurements of  $\overline{\delta V}$  are consistent with the turbulent diffusion rates. Beyond this, the results confirm the previously inferred structure of the wake, but are not of immediate relevance for the present problem of the separation of the diffusive and convective components of the total transport.

An alternative approach may be made by assuming that the eddy diffusion coefficient within the jets is constant, and then choosing values so that the magnitude of jet convection for heat and energy are concordant. The validity of this process is rather debatable, but it can be applied to give consistent results (Fig. 10) by choosing suitable values of  $\kappa_H$ , the diffusion coefficient for heat, and  $\kappa_E$ , the diffusion coefficient for turbulent energy. It is interesting that the value of  $\kappa_H$  found in this way, viz.  $0.018 U_0 d$ , is not far from experimental values for the eddy viscosity, computed in the way described below or directly from the measured shear. The diffusion coefficient for energy is  $0.011 U_0 d$ . This process is rather arbitrary, and results from it may be considerably in error, so these values of the diffusion coefficients should be treated with reserve.

## VI. THE DISTRIBUTION OF MEAN VELOCITY AND TEMPERATURE IN THE WAKE

The distribution of mean velocity in the cylinder wake may be predicted using the methods of the mixing-length theory (14), but the agreement between the theory

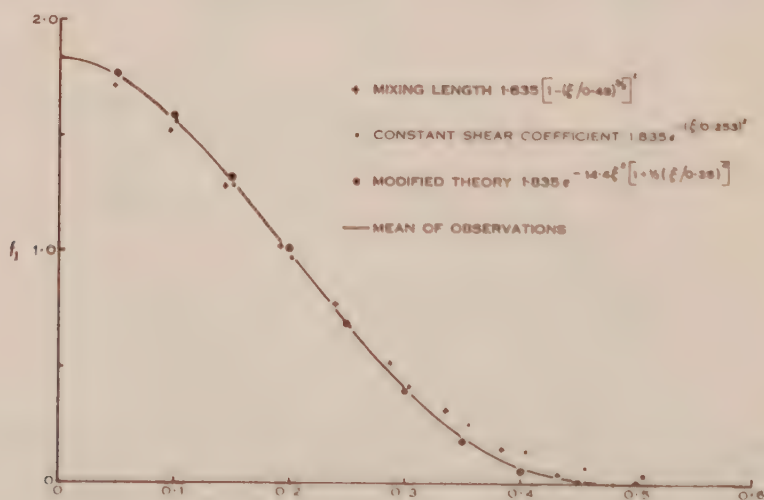


Fig. 11.—Mean velocity distribution.

and experiment, though close, is definitely imperfect, the predicted form of the velocity distribution being too pointed near the maximum and approaching the axis too abruptly (Fig. 11). The hypothesis of constant shear coefficient leads to a better fit over the central portion of the wake, but to a worse fit near the edges. If this hypothesis is modified by assuming that the shear coefficient of the turbulent fluid only is a constant in the wake, and that the non-turbulent fluid between suc-



cessive jets is constrained by pressure gradients to move at the same mean velocity as the fluid in the adjacent jets, the mean effective eddy viscosity is  $\varepsilon/\gamma$ . Then, neglecting the ordinary viscosity  $\nu$  in comparison with  $\varepsilon$ , the non-dimensional velocity distribution function is given by

$$\log_e f_1 = A - \frac{1}{4} \frac{U_0 d}{\varepsilon} \int_0^{\xi} \frac{d\xi^2}{\gamma}.$$

The experimental results for  $\gamma$  may be approximated by

$$\gamma = (1 + \xi^4/\xi_0^4)^{-1},$$

where  $\xi_0 = 0.35$  (see Fig. 6), and then

$$f_1 = A' e^{-\frac{1}{4} \frac{U_0 d}{\varepsilon} \xi^2 \left(1 + \frac{1}{4} \left(\frac{\xi}{0.35}\right)^4\right)}.$$

Selecting values of  $A' = 1.835$  and  $\varepsilon = 0.0173 U_0 d$ ,

$$f_1 = 1.835 e^{-14.4 \xi^2 \left(1 + \frac{1}{4} \left(\frac{\xi}{0.35}\right)^4\right)}$$

which is also plotted in Figure 11. The agreement is good everywhere, and, while the significance of agreement between theoretical predictions and measurements of mean flow distributions is always open to question, the result is useful confirmation of the hypothesis of constant jet eddy viscosity.

Since an appreciable part of the heat transfer is due to jet convection, a similar deduction of the form of the temperature distribution in the wake is not possible, and assumptions must be made concerning the magnitude of the jet convection. Evidently the results could be fitted by choosing an appropriate form for the convection term, but at present this would represent a mathematical exercise without any physical significance. However, any plausible form for the convection term leads to the result that the width of the temperature distribution will be greater than that of the mean velocity distribution.

## VII. GENERAL DISCUSSION

These results are consistent with a wake structure substantially the same as that previously described by the author, but closer study of the flow shows that the small-scale turbulent diffusion of energy and heat down intensity gradients cannot be neglected in comparison with bulk convection due to the large-scale jet motion, and it is necessary to consider the total transport as made up of the sum of gradient diffusion and jet convection. On examining the consequences of this assumption, it is found that a consistent description may be obtained by taking the eddy diffusion coefficients for momentum, energy, and heat as  $0.0173 U_0 d$ ,  $0.011 U_0 d$ , and  $0.018 U_0 d$  respectively. The approximate equality of the diffusion coefficients for momentum and heat indicates that the same processes are responsible for the gradient diffusion of both quantities, and that the striking differences between the total

transport rates are caused by the large-scale jet convection becoming comparable with the small-scale diffusive transport. To understand this, it must be remembered that, while heat can only be transferred by diffusive movements of the fluid, transfer of momentum also takes place by the action of pressure gradients. Thus momentum transfer in the  $Oy$  direction takes place inside the jets by diffusive movements, and then in the  $Ox$  direction to the adjoining non-turbulent fluid by longitudinal pressure gradients. If the longitudinal transfer is efficient, then the mean velocity will be much the same at any point, regardless of whether or not the fluid is turbulent at the time of observation, and large-scale fluctuations of the down-stream velocity component will be rare. Measurements of the down-stream scale,

$$L_u = U_0 \int_0^{\infty} \frac{\overline{u(t)u(t+r)}}{\overline{u^2}} dr,$$

show that such fluctuations are not very common, and are no more frequent in the regions of maximum shear than elsewhere, and it seems probable that the mean velocity gradient in the jets is not significantly different from the gradient in the surrounding non-turbulent fluid. On the other hand, the temperature in the non-turbulent fluid must be the ambient temperature, at least at high Reynolds numbers, and since the mean temperature within the jets varies relatively little with position, the gradient of the mean local temperature is small, and diffusive transport is correspondingly small and comparable in magnitude with transport by jet convection. It is this change in the relative proportions of the two components of the transfer that gives rise to the apparently greater rate of heat diffusion. Many of the same considerations apply to the transfer of turbulent energy from one part of the wake to another, but the problem is complicated by the dissipation and pressure flow of turbulent energy.

The existence of eddy diffusion coefficients for momentum, turbulent energy, and heat seems to show that the diffusive processes take place in times short compared with  $x/U_0$ , the time necessary for development of the wake, and over distances small compared with  $\sqrt{(x-x_0)d}$ , the width of the wake. It is easily verified that this is true, writing the eddy diffusion coefficient as the product of a mixing velocity and a mixing scale, as is usual in mixing-length theory. Using the experimental value of  $\sqrt{v^2}$  as the mixing velocity, viz.  $0.3 U_0 \sqrt{d/(x-x_0)}$ , and taking the eddy diffusion coefficient to be  $0.017 U_0 d$ , the mixing scale is  $0.05 \sqrt{(x-x_0)d}$ , and the time of mixing is  $0.19 (x-x_0)/U_0$ , both of which are reasonably small. In this way, by considering the experimental data on diffusion and turbulent intensities, the contradiction in mixing-length theory, that the processes could not be on a sufficiently small scale, is removed, and a local transfer coefficient is shown to exist, although its value is determined by the overall and not the local characteristics of the wake.

Within experimental error, the eddy viscosity is a constant over the whole wake, and it is likely that this is also true for the eddy conductivity. This is consistent with observations that length parameters such as  $\lambda_u$  or  $L_u$  and the turbulent energy do not vary greatly across the wake, and the small-scale turbulent motion within the jets must be surprisingly homogeneous. It has frequently been noted that  $\lambda_u$  is nearly constant in a region of shear flow, and this is a little surprising, as there is no apparent reason why  $\lambda_u$ , a length determined by the ratio of the intensity of one velocity component to the total viscous dissipation, should be any less variable than the similar lengths involving the intensities of the other two components. The present results (Fig. 7) show that the variations of  $\lambda_u$  are considerable, and that it is the local viscous dissipation

$$15 \frac{\nu}{\gamma} \overline{\left( \frac{\partial u}{\partial x} \right)^2}$$

which is more nearly constant (Fig. 12). At higher Reynolds numbers the variation

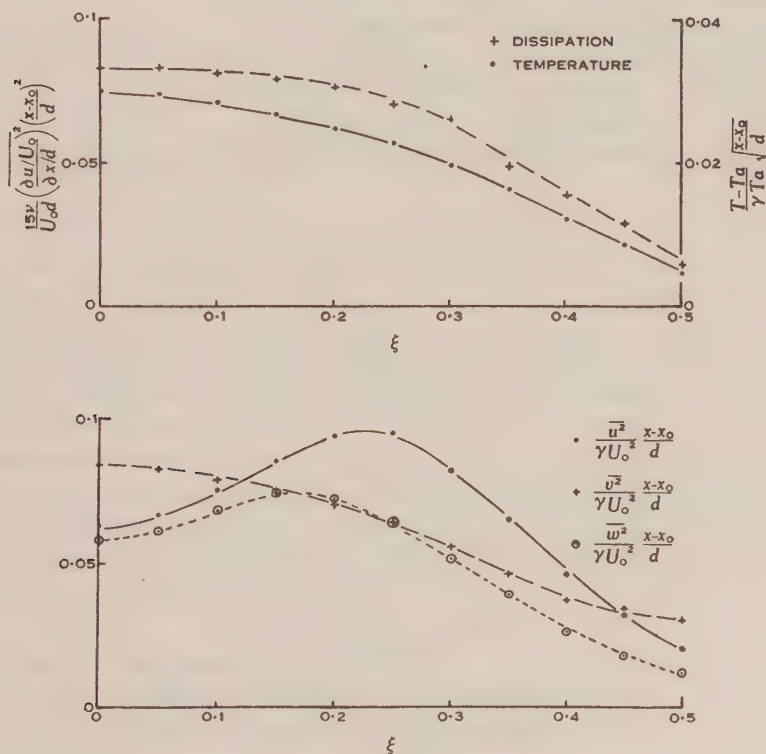


Fig. 12.—Local properties of the jets.

is even less, and there seems to be a strong tendency toward spatial homogeneity of  $\overline{(\partial u / \partial x)^2}$ . This tendency has also been found by R. Stewart (personal communica-

tion) who investigated the mixing zone behind a grid of parallel, equally spaced, circular cylinders, and found that spatial variations in the intensities of velocity components persist after spatial variations in the intensities of velocity derivatives have disappeared. Accepting the substantial constancy of  $(\partial u / \partial x)^2$  across the wake, the approximate constancy of  $\lambda_u$  is merely a consequence of the similar behaviour of  $\overline{u^2}/\gamma$  and the more precise constancy of the viscous dissipation.

### VIII. REFERENCES

- (1) CHOU, P. Y.—*Quart. Appl. Math.* **3**: 38-54 (1945).
- (2) CHOU, P. Y.—*Ibid.* **3**: 198-209 (1945).
- (3) HU, N.—*Chin. J. Phys.* **5**: 30 (1944).
- (4) CORRSIN, S.—U.S. Nat. Adv. Ctee. Aero. War Rep. W-94 (1943).
- (5) CORRSIN, S., and UBEROI, M. S.—Guggenheim Aero. Lab. Pasadena, Rep. No. 3661 (1947).
- (8) DRYDEN, H.—*Advances Appl. Mech.* **1**: 1-40 (1948).
- (9) TOWNSEND, A. A.—*Proc. Roy. Soc. A* **190**: 551-61 (1947).
- (10) TOWNSEND, A. A.—*Ibid.* **197**: 124-40 (1949).
- (11) TOWNSEND, A. A.—*Proc. Camb. Phil. Soc.* **43**: 560-70 (1947).
- (12) TOWNSEND, A. A.—*Aust. J. Sci. Res. A* **1**: 161-74 (1948).
- (13) CORRSIN, S.—*Rev. Sci. Instrum.* **18**: 469-71 (1947).
- (14) GOLDSTEIN, S.—“Modern Developments in Fluid Dynamics.” (Clarendon Press : Oxford, 1938.)



# THE UNIFORM FLEXURE OF AN INCOMPLETE TORE

By W. FREIBERGER\* and R. C. T. SMITH†

[Manuscript received May 6, 1949]

## Summary

In this paper we discuss the flexure of an incomplete tore in the plane of its circular centre-line. We reduce the problem to the determination of two harmonic functions, subject to boundary conditions on the surface of the tore which involve the first two derivatives of the functions. We point out the relation of this solution to the general solution of three-dimensional elasticity problems.

The special case of a narrow rectangular cross-section is solved exactly in Appendix II.

## I. INTRODUCTION

Michell(1) and Southwell(2) reduced the solution of the flexure problem for incomplete tores to the determination of a single stress function which satisfies an equation of the form

$$\frac{\partial^4 \Psi}{\partial r^4} + 2 \frac{\partial^4 \Psi}{\partial r^2 \partial z^2} + \frac{\partial^4 \Psi}{\partial z^4} - \frac{2}{r} \frac{\partial^3 \Psi}{\partial r^3} - \frac{2}{r} \frac{\partial^3 \Psi}{\partial r \partial z^2} + \frac{3}{r^2} \frac{\partial^2 \Psi}{\partial r^2} - \frac{3}{r} \frac{\partial \Psi}{\partial r} = \text{constant};$$

the boundary conditions to be satisfied on the lateral surface of the tore consist of two equations each of which involves derivatives of  $\Psi$  up to the third order. Michell started from expressions for the displacements derived from symmetry considerations. Southwell used equilibrium equations for the stresses in conjunction with the Beltrami-Michell compatibility relations, without explicit reference to strains. Equilibrium equations in terms of displacements were derived by Fubini(3). Tricomi(4) has used these equations in a discussion of the rectangular cross-section.

The essential point in the present solution is the realization that the displacement which gives rise to the dilatation is the gradient of a scalar potential function  $F$  and the displacement which produces the rotation is the curl of a solenoidal vector potential function  $G$  ("Helmholtz transformation," see Mindlin 5). The displacement vector  $\mathbf{u}$  can therefore be written

$$\mathbf{u} = \text{grad } F + \text{curl } G$$

and taking the divergence of both sides

$$\text{div } \mathbf{u} = \text{div grad } F, \text{ since } \text{div curl} \equiv 0.$$

But  $\text{div } \mathbf{u} = \theta$ , the cubical dilatation; consequently

$$\theta = \nabla^2 F.$$

\*Aeronautical Research Laboratories, Department of Supply and Development, Melbourne.

†Department of Mathematics and Physics, New England University College, Armidale, N.S.W. Formerly: Aeronautical Research Laboratories, Department of Supply and Development, Melbourne.

In the following solution the function  $F$  is introduced in the Navier equations of equilibrium by means of this substitution, after the terms containing the angular displacement have been eliminated by another substitution.

Finally the solution is reduced to the determination of two functions  $\varphi_0$  and  $\varphi_1$ , each satisfying Laplace's equation

$$\nabla^2 \varphi = \frac{\partial^2 \varphi}{\partial r^2} + \frac{\partial^2 \varphi}{\partial z^2} + \frac{1}{r} \frac{\partial \varphi}{\partial r} = 0,$$

the two boundary conditions involving only first and second derivatives of  $\varphi_0$ ,  $\varphi_1$ . The solution for the narrow rectangular cross-section is completed in Appendix II.

The method is similar to that of Papkovitch-Neuber(6) for the solution of three-dimensional problems. In the latter, four potential functions enter naturally into the solution and it is shown that this number can be reduced to three. In the present paper a solution originally involving two potential functions and a third function satisfying  $(\nabla^2 - 1/r^2)\varphi = 0$  is expressed in terms of two harmonic functions. It is not a special case of the Papkovitch-Neuber solution since the latter is only valid for rectangular Cartesian coordinates. (The Papkovitch-Neuber solution can be generalized to arbitrary coordinates but the treatment is elaborate and the results complicated; see Freiburger 7). The problem is virtually two-dimensional and it is perhaps of interest that a pair of equations analogous to the Cauchy-Riemann relations in the theory of functions is found. Such pairs of equations also occur in hydrodynamics and have been discussed by Bers and Gelbart(8).

In conclusion, an expression for the resultant couple is obtained which is simpler than that in previous solutions.

## II. SOLUTION OF THE PROBLEM

### (a) *Solution Associated with the Angular Displacement*

The fundamental equations in cylindrical coordinates are given in various books, e.g. Sokolnikoff(9).

Assuming the stresses are independent of  $\delta$  we find for the equations of equilibrium in terms of displacements

$$\left. \begin{aligned} \frac{\partial^2 u}{\partial z^2} - \frac{\partial^2 w}{\partial r \partial z} - \frac{1}{r^2} \frac{\partial v}{\partial \delta} - \frac{1}{r} \frac{\partial^2 v}{\partial r \partial \delta} + \frac{\lambda + 2\mu}{\mu} \frac{\partial \theta}{\partial r} &= 0, \\ \frac{\partial^2 v}{\partial r^2} + \frac{1}{r} \frac{\partial v}{\partial r} + \frac{\partial^2 v}{\partial z^2} - \frac{1}{r^2} v &= 0, \\ \frac{\partial^2 w}{\partial r^2} + \frac{1}{r} \frac{\partial w}{\partial r} - \frac{1}{r} \frac{\partial u}{\partial z} - \frac{\partial^2 u}{\partial r \partial z} - \frac{1}{r} \frac{\partial^2 v}{\partial z \partial \delta} + \frac{\lambda + 2\mu}{\mu} \frac{\partial \theta}{\partial z} &= 0, \end{aligned} \right\} \dots(1)$$

where

$$\theta = \frac{1}{r} \frac{\partial(ru)}{\partial r} + \frac{1}{r} \frac{\partial v}{\partial \delta} + \frac{\partial w}{\partial z} \dots\dots\dots(2)$$

is the dilatation.

The immediate aim is to find a simple basic system of displacements satisfying (1).

For flexure the meridian sections remain plane and become meridian sections in the deformed tore which will be supposed coaxial with its original form. The angular displacement must therefore have the form

$$v' = Kr\delta. \quad \dots\dots\dots(3)$$

It is easy to verify that the system

$$u' = \frac{\mu K}{\lambda + 2\mu} r \log r, \quad v' = Kr\delta, \quad w' = 0, \quad \dots\dots\dots(4)$$

satisfies the equilibrium equations (1).

### (b) The Complete Solution

To make up the general system we will superimpose the unknown displacements  $u''$ ,  $w''$  on the system (4).

$$u = \frac{\mu K}{\lambda + 2\mu} r \log r + u'', \quad v = Kr\delta, \quad w = w'' \quad \dots\dots\dots(5)$$

Substituting the system (5) in (1) gives

$$\left. \begin{aligned} \frac{\partial^2 u''}{\partial z^2} - \frac{\partial^2 w''}{\partial r \partial z} + \frac{\lambda + 2\mu}{\mu} \frac{\partial \theta''}{\partial r} &= 0, \\ \frac{\partial^2 w''}{\partial r^2} + \frac{1}{r} \frac{\partial w''}{\partial r} - \frac{1}{r} \frac{\partial u''}{\partial r} - \frac{\partial^2 u''}{\partial r \partial z} + \frac{\lambda + 2\mu}{\mu} \frac{\partial \theta''}{\partial r} &= 0, \end{aligned} \right\} \dots\dots\dots(6)$$

where

$$\theta'' = \frac{\partial u''}{\partial r} + \frac{\partial w''}{\partial z} + \frac{u''}{r}. \quad \dots\dots\dots(7)$$

The second equation of (1) is satisfied identically.

It is convenient to drop accents from now on. Simplifying (6) gives

$$\left. \begin{aligned} \left( \nabla^2 - \frac{1}{r^2} \right) u + \frac{\lambda + \mu}{\mu} \frac{\partial \theta}{\partial r} &= 0, \\ \nabla^2 w + \frac{\lambda + \mu}{\mu} \frac{\partial \theta}{\partial z} &= 0, \end{aligned} \right\} \dots\dots\dots(8)$$

where

$$\nabla^2 \equiv \frac{\partial^2}{\partial r^2} + \frac{1}{r} \frac{\partial}{\partial r} + \frac{\partial^2}{\partial z^2}$$

is the Laplacian operator. Equations (8) can be deduced from those of Fubini(3) by a substitution equivalent to the first of system (5).

As explained in the introduction the substitution

$$\theta = \nabla^2 F_0 \quad \dots\dots\dots(9)$$

will be used in the solution of (8). Equations (8) become

$$\left. \begin{aligned} \left( \nabla^2 - \frac{1}{r^2} \right) \left( u + \frac{\lambda + \mu}{\mu} \frac{\partial F_0}{\partial r} \right) &= 0 \\ \nabla^2 \left( w + \frac{\lambda + \mu}{\mu} \frac{\partial F_0}{\partial z} \right) &= 0 \end{aligned} \right\} \dots\dots\dots(10)$$

on simplifying by means of the relations

$$\frac{\partial}{\partial r} \nabla^2 = \left( \nabla^2 - \frac{1}{r^2} \right) \frac{\partial}{\partial r}, \quad \frac{\partial}{\partial z} \nabla^2 = \nabla^2 \frac{\partial}{\partial z} \quad \dots\dots\dots(11)$$

The solution of (10) is

$$u = -\frac{\lambda + \mu}{\mu} \frac{\partial F_0}{\partial r} + u_0, \quad w = -\frac{\lambda + \mu}{\mu} \frac{\partial F_0}{\partial z} + w_1, \quad \dots\dots(12)$$

where

$$\left( \nabla^2 - \frac{1}{r^2} \right) u_0 = 0, \quad \nabla^2 w_1 = 0, \quad \dots\dots\dots(13)$$

the choice of indices being made for later convenience. Next, the form of  $F$  will be determined. Substituting (9) and (12) in (7) we obtain

$$\frac{\lambda + 2\mu}{\mu} \nabla^2 F_0 = \frac{\partial w_1}{\partial z} + \frac{\partial u_0}{\partial r} + \frac{u_0}{r}, \quad \dots\dots\dots(14)$$

which has the solution

$$F_0 = \frac{\mu}{2(\lambda + 2\mu)} (ru_0 + zw_1) + F_1, \quad \dots\dots\dots(15)$$

where

$$\nabla^2 F_1 = 0. \quad \dots\dots\dots(16)$$

The solution of the equilibrium equations (8) is given by (12) on substituting (15).

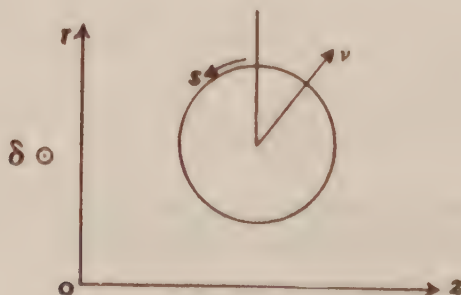


Fig. 1



## III. REDUCTION OF THE SOLUTION

The function  $u_0$  will now be eliminated from the solution of the previous section.

A function  $w_0$  "conjugate" to  $u_0$  is defined by the equations

$$\frac{\partial w_0}{\partial z} = \left( \frac{\partial}{\partial r} + \frac{1}{r} \right) u_0, \quad \frac{\partial w_0}{\partial r} = - \frac{\partial u_0}{\partial z}. \quad (17)$$

The condition for integrability of these equations is simply

$$(\nabla^2 - 1/r^2)u_0 = 0, \quad (18)$$

and on eliminating  $u_0$ ,  $w_0$  is found to satisfy

$$\nabla^2 w_0 = 0. \quad (19)$$

Equations (17) are analogous to the Cauchy-Riemann relations.

A remarkable result is that if  $u_0$ ,  $w_0$  are conjugates, then

$$\left( \nabla^2 - \frac{1}{r^2} \right) (zu_0 + rw_0) = 0 \quad (20)$$

$$\nabla^2 (ru_0 - zw_0) = 0. \quad (21)$$

However,  $(zu_0 + rw_0)$ ,  $(ru_0 - zw_0)$  are not conjugates.

The result (21) will now be used to simplify the expressions for  $u$ ,  $w$ . First consider  $F$ .

From (21)

$$ru_0 = zw_0 + F_2, \quad (22)$$

where  $F_2$  is harmonic. Then (15) becomes

$$F_0 = \frac{\mu}{2(\lambda + 2\mu)} z(w_0 + w_1) + F_1 + F_2 \quad (23)$$

On writing

$$u_0 = - \frac{\lambda + \mu}{\mu} \frac{\partial F_3}{\partial r}, \quad (24)$$

$$F^* = \frac{\mu}{2(\lambda + 2\mu)} z(w_0 + w_1) + F_1 + F_2 + F_3, \quad (25)$$

equation (12) becomes

$$u = - \frac{\lambda + \mu}{\mu} \frac{\partial F^*}{\partial r}, \quad w = - \frac{\lambda + \mu}{\mu} \frac{\partial F^*}{\partial z} + w_1 - \frac{\lambda + \mu}{\mu} \frac{\partial F_3}{\partial z} \quad (26)$$

It is straightforward to show that  $F_3$  can be chosen as a harmonic function. For

$$\left( \nabla^2 - \frac{1}{r^2} \right) u_0 = 0 = \left( \nabla^2 - \frac{1}{r^2} \right) \frac{\partial F_3}{\partial r} = \frac{\partial}{\partial r} \nabla^2 F_3,$$

that is

$$\nabla^2 F_3 = f(z),$$

and  $F_3$  can be made harmonic by adding to it a suitable function of  $z$ , which of course does not affect  $\partial F_3 / \partial r$ .

Now from (24) and (17)

$$-\frac{\lambda + \mu}{\mu} \frac{\partial^2 F_3}{\partial r \partial z} = \frac{\partial u_0}{\partial z} = \frac{\partial w_0}{\partial r},$$

and

$$-\frac{\lambda + \mu}{\mu} \frac{\partial F_3}{\partial z} = w_0, \quad \dots\dots\dots(27)$$

where the arbitrary function of  $z$  introduced by integration with respect to  $r$  is necessarily linear, since  $w_0$ ,  $\partial F_3/\partial z$  are harmonic and can obviously be absorbed in  $F_3$ . Therefore

$$u = -\frac{\lambda + \mu}{\mu} \frac{\partial F^*}{\partial r}, \quad w = -\frac{\lambda + \mu}{\mu} \frac{\partial F^*}{\partial z} + w_0 + w_1. \quad \dots(28)$$

To simplify the notation put

$$\left. \begin{aligned} \varphi_0 &= w_0 + w_1, \\ \varphi_1 &= \frac{\lambda + \mu}{\mu} (F_1 + F_2 + F_3), \\ F &= \frac{\lambda + \mu}{2(\lambda + 2\mu)} z \varphi_0 + \varphi_1. \end{aligned} \right\} \quad \dots\dots\dots(29)$$

Then (28) becomes

$$u = -\frac{\partial F}{\partial r}, \quad w = -\frac{\partial F}{\partial z} + \varphi_0, \quad \dots\dots\dots(30)$$

where  $\varphi_0$ ,  $F$  are given by (29) and  $\varphi_0$ ,  $\varphi_1$  are harmonic.

By a quite similar argument it would have been possible to simplify (12) to the form

$$u = -\frac{\partial F}{\partial r} + \Psi_0, \quad w = -\frac{\partial F}{\partial z}, \quad \dots\dots\dots(31)$$

$$F = \frac{\lambda + \mu}{2(\lambda + 2\mu)} r \Psi_0 + \Psi_1,$$

where

$$(\nabla^2 - 1/r^2) \Psi_0 = 0, \quad \nabla^2 \Psi_1 = 0.$$

Since  $\Psi_0$ ,  $\Psi_1$  satisfy different differential equations this is not as convenient as (30).

#### IV. THE STRESSES

The stresses corresponding to the displacements found in the previous paragraphs will now be evaluated.

The strains corresponding to the basic system of displacements (4) are

$$\left. \begin{aligned} e'_{rr} &= \frac{\mu K}{\lambda + 2\mu} (\log r + 1), \quad e'_{\delta\delta} = \frac{\mu K}{\lambda + 2\mu} \log r + K \\ e'_{rz} &= e'_{r\delta} = e'_{rz} = e'_{\delta z} = 0; \end{aligned} \right\} \quad \dots(32)$$

and the stresses

$$\left. \begin{aligned} \widehat{rr}' &= \frac{2\mu K}{\lambda + 2\mu} (\lambda + \mu) \log r + K(\lambda + \mu), \\ \widehat{\delta\delta}' &= \frac{2\mu K}{\lambda + 2\mu} (\lambda + \mu) \log r + \frac{\lambda\mu K}{\lambda + 2\mu} + K(\lambda + 2\mu), \\ \widehat{zz}' &= \frac{2\lambda\mu K}{\lambda + 2\mu} \log r + \frac{\lambda\mu K}{\lambda + 2\mu} + \lambda K, \\ \widehat{r\delta}' &= \widehat{z\delta}' = \widehat{rz}' = 0. \end{aligned} \right\} \dots(33)$$

The displacements (30) correspond to the following strains and stresses :

$$\left. \begin{aligned} e_{rr} &= -\frac{\partial^2 F}{\partial r^2}, \quad e_{\delta\delta} = -\frac{1}{r} \frac{\partial F}{\partial r}, \quad e_{zz} = -\frac{\partial^2 F}{\partial z^2} + \frac{\partial \varphi_0}{\partial z} \\ e_{r\delta} &= 0, \quad e_{\delta z} = 0, \quad e_{rz} = \frac{1}{2} \left( \frac{\partial \varphi_0}{\partial r} + 2 \frac{\partial^2 F}{\partial r \partial z} \right) \end{aligned} \right\} \dots(34)$$

therefore

$$\left. \begin{aligned} \theta &= e_{rr} + e_{\delta\delta} + e_{zz} = -\nabla^2 F + \frac{\partial \varphi_0}{\partial z} \\ \widehat{rr} &= \lambda \left( -\nabla^2 F + \frac{\partial \varphi_0}{\partial z} \right) - 2\mu \frac{\partial^2 F}{\partial r^2}, \\ \widehat{\delta\delta} &= \lambda \left( -\nabla^2 F + \frac{\partial \varphi_0}{\partial z} \right) - \frac{2\mu}{r} \frac{\partial F}{\partial r}, \\ \widehat{zz} &= \lambda \left( -\nabla^2 F + \frac{\partial \varphi_0}{\partial z} \right) + 2\mu \left( -\frac{\partial^2 F}{\partial z^2} + \frac{\partial \varphi_0}{\partial z} \right), \\ \widehat{rz} &= -2\mu \frac{\partial^2 F}{\partial r \partial z} + \mu \frac{\partial \varphi_0}{\partial r}, \quad \widehat{r\delta} = \widehat{\delta z} = 0. \end{aligned} \right\} \dots(35)$$

It follows from (33) and (35) that there are no tangential stresses acting on axial planes, because

$$\widehat{r\delta}' = r\delta = \widehat{\delta z}' = \widehat{\delta z} = 0.$$

To express (35) in terms of  $\varphi_0$  and  $\varphi_1$ , the last of equations (29) will be substituted for  $F$ ; the result is

$$\left. \begin{aligned} \widehat{rr} &= \frac{\lambda\mu}{\lambda + 2\mu} \frac{\partial \varphi_0}{\partial z} - \mu \frac{\lambda + \mu}{\lambda + 2\mu} z \frac{\partial^2 \varphi_0}{\partial r^2} - 2\mu \frac{\partial^2 \varphi_1}{\partial r^2}, \\ \widehat{\delta\delta} &= \frac{\lambda\mu}{\lambda + 2\mu} \frac{\partial \varphi_0}{\partial z} - \mu \frac{\lambda + \mu}{\lambda + 2\mu} \frac{z}{r} \frac{\partial \varphi_0}{\partial r} - \frac{2\mu}{r} \frac{\partial \varphi_1}{\partial r}, \\ \widehat{zz} &= \mu \frac{\partial \varphi_0}{\partial z} - \mu \frac{\lambda + \mu}{\lambda + 2\mu} z \frac{\partial^2 \varphi_0}{\partial z^2} - 2\mu \frac{\partial^2 \varphi_1}{\partial z^2}, \\ \widehat{rz} &= \frac{\mu^2}{\lambda + 2\mu} \frac{\partial \varphi_0}{\partial r} - \mu \frac{\lambda + \mu}{\lambda + 2\mu} z \frac{\partial^2 \varphi_0}{\partial r \partial z} - 2\mu \frac{\partial^2 \varphi_1}{\partial r \partial z}, \\ \widehat{r\delta} &= \widehat{\delta z} = 0. \end{aligned} \right\} \dots(36)$$

## V. BOUNDARY CONDITIONS

The boundary traction on any point of the lateral surface of the tore is specified by its components  $\widehat{rv}$ ,  $\widehat{vs}$ , and  $\widehat{v\delta}$ ;  $\delta$  denoting a direction perpendicular to the axial

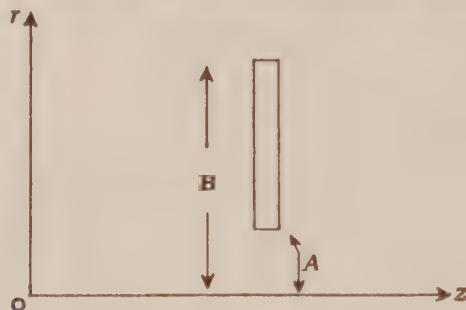


Fig. 2

plane,  $v$  the outward-drawn normal, and  $s$  being measured along a generating curve, the following transformation equations hold (see Southwell(2)):

$$\widehat{vr} = \widehat{rr} \cos(r, v) + \widehat{zr} \sin(r, v) \quad \dots\dots\dots(37)$$

$$\widehat{v\delta} = \widehat{r\delta} \cos(r, v) + \widehat{\delta z} \sin(r, v) \quad \dots\dots\dots(38)$$

$$\widehat{vz} = \widehat{zr} \cos(r, v) + \widehat{zz} \sin(r, v) \quad \dots\dots\dots(39)$$

From the last of (36) and from (38) it follows that  $\widehat{v\delta}$  vanishes identically. The lateral surface of the tore is stress-free, so that (37) and (39) give the two boundary conditions

$$\left. \begin{aligned} \widehat{vr} &= \widehat{rr} \cos(r, v) + \widehat{zr} \sin(r, v) = 0, \\ \widehat{vz} &= \widehat{zr} \cos(r, v) + \widehat{zz} \sin(r, v) = 0, \end{aligned} \right\} \quad \dots\dots\dots(40)$$

on the boundary  $C$ .

It must be noted that the stresses in (40) include both the basic system (33) and the stresses given by (36).

$$\text{Since } \cos(r, v) = \frac{\partial r}{\partial v}, \quad \sin(r, v) = \frac{\partial z}{\partial v}$$

and

$$\frac{\partial r}{\partial v} = -\frac{\partial z}{\partial s}, \quad \frac{\partial z}{\partial v} = \frac{\partial r}{\partial s}$$

substituting (36) in (40) gives the two equations to be satisfied on the lateral surface of the tore

$$\begin{aligned} \frac{\mu}{\lambda + 2\mu} \left( \lambda \frac{\partial q_0}{\partial z} \frac{\partial r}{\partial v} + \mu \frac{\partial q_0}{\partial r} \frac{\partial z}{\partial v} \right) - \mu \frac{\lambda + \mu}{\lambda + 2\mu} z \frac{\partial}{\partial v} \left( \frac{\partial q_0}{\partial r} \right) - 2\mu \frac{\partial}{\partial v} \left( \frac{\partial q_1}{\partial r} \right) \\ + \frac{\partial r}{\partial v} K(\lambda + \mu) \left( \frac{2\mu}{\lambda + 2\mu} \log r + 1 \right) = 0, \quad \dots\dots\dots(41) \end{aligned}$$



$$\begin{aligned} \mu \left( \frac{\mu}{\lambda + 2\mu} \frac{\partial \varphi_0}{\partial r} \frac{\partial r}{\partial v} + \frac{\partial \varphi_0}{\partial z} \frac{\partial z}{\partial v} \right) - \mu \frac{\lambda + \mu}{\lambda + 2\mu} z \frac{\partial}{\partial v} \left( \frac{\partial \varphi_0}{\partial z} \right) - 2\mu \frac{\partial}{\partial v} \left( \frac{\partial \varphi_1}{\partial z} \right) \\ + \frac{\partial z}{\partial v} K \lambda \left( \frac{2\mu}{\lambda + 2\mu} \log r + \frac{\mu}{\lambda + 2\mu} + 1 \right) = 0 \quad \dots\dots(42) \end{aligned}$$

## VI. RESULTANT FORCES

It will now be shown that the resultant of the stresses over any cross-section is the required flexural couple.

Since  $\widehat{r\delta} = \widehat{\delta z} = 0$  identically throughout the tore, the only stress acting on a cross-section is  $\widehat{\delta\delta}$ ; its resultant:

$$\begin{aligned} \iint \widehat{\delta\delta} \, dr \, dz &= \iint \left\{ \frac{\partial}{\partial r} (r \cdot \widehat{rr}) + \frac{\partial}{\partial z} (r \cdot \widehat{zr}) \right\} dr \, dz, \\ &= \int_C r \left\{ \widehat{rr} \cos(r, v) + \widehat{zr} \sin(r, v) \right\} ds \text{ by Stokes's theorem,} \\ &= 0 \quad \text{by the first of (40).} \end{aligned}$$

The resultant couple about a radius  $Or$ :

$$\begin{aligned} \iint z \cdot \widehat{\delta\delta} \, dr \, dz &= \iint \left\{ \frac{\partial}{\partial r} (zr \cdot \widehat{rr} - r^2 \cdot \widehat{zr}) + \frac{\partial}{\partial z} (zr \cdot \widehat{zr} - r^2 \cdot \widehat{zz}) \right\} dr \, dz, \\ &= \int_C \{ zr [\widehat{rr} \cos(r, v) + \widehat{zr} \sin(r, v)] \\ &\quad - r^2 [\widehat{zr} \cos(r, v) + \widehat{zz} \sin(r, v)] \} ds \text{ by Stokes's theorem,} \\ &= 0 \quad \text{by (40).} \end{aligned}$$

The only non-zero resultant is therefore a couple with axis along  $Oz$ , as required. Its magnitude is

$$\begin{aligned} \iint r \cdot \widehat{\delta\delta} \, dr \, dz &= \iint \left\{ \frac{\lambda\mu}{\lambda + 2\mu} r \frac{\partial \varphi_0}{\partial z} - \mu \frac{\lambda + \mu}{\lambda + 2\mu} z \frac{\partial \varphi_0}{\partial r} - 2\mu \frac{\partial \varphi_1}{\partial r} \right\} dr \, dz \\ &\quad + \frac{2K\mu}{\lambda + 2\mu} (\lambda + \mu) \iint r \log r \, dr \, dz + K \left( \frac{\lambda\mu}{\lambda + 2\mu} + \lambda + 2\mu \right) \iint r \, dr \, dz \end{aligned}$$

by (33) and (36).

Now,

$$\begin{aligned} \iint \left\{ \frac{\lambda\mu}{\lambda + 2\mu} r \frac{\partial \varphi_0}{\partial z} - \mu \frac{\lambda + \mu}{\lambda + 2\mu} z \frac{\partial \varphi_0}{\partial r} \right\} dr \, dz \\ = \frac{1}{\lambda + 2\mu} \iint \left\{ \frac{\partial}{\partial z} [\lambda \mu r \varphi_0] + \frac{\partial}{\partial r} [-\mu (\lambda + \mu) z \varphi_0] \right\} dr \, dz, \\ = \frac{1}{\lambda + 2\mu} \int_C \left\{ \varphi_0 [-\mu (\lambda + \mu) z \cos(r, v) + \lambda \mu r \sin(r, v)] \right\} ds \\ \hspace{15em} \text{by Stokes's theorem.} \end{aligned}$$

Hence the resultant couple becomes

$$\begin{aligned}
 R &= \iint r \delta \delta \, dr dz \\
 &= \frac{1}{\lambda + 2\mu} \int_C \left\{ \varphi_0 \left[ -\mu(\lambda + \mu) + z \frac{\partial r}{\partial v} + \lambda \mu r \frac{\partial z}{\partial v} \right] \right\} ds \\
 &\quad - 2\mu \int_D \frac{\partial \varphi_1}{\partial r} \, dr dz + \frac{2K\mu}{\lambda + 2\mu} (\lambda + \mu) \iint r \log r \, dr dz \\
 &\quad + K \left( \frac{\lambda \mu}{\lambda + 2\mu} + \lambda + 2\mu \right) \iint r \, dr dz, \quad \dots\dots\dots (43)
 \end{aligned}$$

where  $C$  is the boundary of any cross-section and  $D$  its area.

## VII. CONCLUSION

The deformation by a flexural couple of a bar curved into the form of a circle (an incomplete tore) is considered in this paper by a stress function method. The stress system corresponds to that set up in a helical spring under axial torsion.

The problem is reduced to the determination of two functions  $\varphi_0$  and  $\varphi_1$  each satisfying Laplace's equation

$$\nabla^2 \varphi = 0$$

at every point of the cross-section and subject to the boundary conditions (41)-(42) on the circumference.

The stress system will then be given by the sum of the two systems (33) and (36), and the resultant flexural couple by (43).

The special case of a narrow rectangular cross-section is solved exactly in Appendix II.

## VIII. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Aeronautical Research Laboratories. The authors are indebted to the Chief Scientist, Department of Supply and Development, for permission to publish this paper.

## IX. REFERENCES

- (1) MICHELL, J. H. —The uniform torsion and flexure of incomplete tores. *Proc. Lond. Math. Soc.* **31**: 130-46 (1899).
- (2) SOUTHWELL, R. V. —Some practically important stress systems in solids of revolution. *Proc. Roy. Soc. A* **180**: 367-96 (1942).
- (3) FUBINI, G. —Sulla teoria della trave ad asse curvo. *Atti. Accad. Sci. Torino, Cl. I* **72**: 386-410 (1937).
- (4) TRICOMI, F. —Su di un notevole caso di deformazione di una trave curva. *Ibid.* **73**: 79-109 (1937).

- (5) MINDLIN, R. D.—Note on the Galerkin and Papkovitch stress functions. *Bull. Amer. Math. Soc.* **42**: 373 (1936).
- (6) NEUBER, H.—Eine neue Methode zur Lösung 3-dimensionalen Elastizitätsprobleme. *Z. Angew. Math. Mech.* **14**: 203 (1934).
- (7) FREIBERGER, W.—On the solution of the equilibrium equations of elasticity in general curvilinear coordinates. *Aust. J. Sci. Res. A.* **2**: 483 (1949).
- (8) BERS, L., and GELBART, A.—On a class of differential equations in mechanics of continua. *Quart. Appl. Math.* **1**: 168 (1943).
- (9) SOKOLNIKOFF, I. S.—“Mathematical Theory of Elasticity.” 1st Ed. p. 202. (McGraw-Hill Book Co.: New York and London, 1946.)
- (10) TIMOSHENKO, S.—“Theory of Elasticity.” 1st Ed. para. 23. (McGraw-Hill Book Co.: New York and London, 1934.)

## APPENDIX I

*Curvilinear Coordinates*

For certain cross-sections of the tore the cylindrical polar coordinates hitherto used will not be the most convenient system. Circular cross-sections, for instance, will suggest the system of toroidal coordinates since this will transform the lateral surface of the tore into a coordinate surface and so simplify the boundary conditions.

The solution of Section III will now be transformed from the  $r, z$  coordinates into arbitrary curvilinear orthogonal coordinates  $\alpha, \beta$ ;  $\delta$  will naturally be retained.

The transformation equations are

$$\frac{u_\alpha}{h_\alpha} = u \frac{\partial \alpha}{\partial r} + w \frac{\partial \alpha}{\partial z}, \quad \frac{u_\beta}{h_\beta} = u \frac{\partial \beta}{\partial r} + w \frac{\partial \beta}{\partial z}, \quad \text{for displacements,}$$

$$\frac{1}{h_\alpha^2} \frac{\partial F}{\partial \alpha} = \frac{\partial F}{\partial r} \frac{\partial \alpha}{\partial r} + \frac{\partial F}{\partial z} \frac{\partial \alpha}{\partial z}, \quad \frac{1}{h_\beta^2} \frac{\partial F}{\partial \beta} = \frac{\partial F}{\partial r} \frac{\partial \beta}{\partial r} + \frac{\partial F}{\partial z} \frac{\partial \beta}{\partial z};$$

substituting from (30)

$$\frac{1}{h_\alpha^2} \frac{\partial F}{\partial \alpha} = -u \frac{\partial \alpha}{\partial r} - w \frac{\partial \alpha}{\partial z} + \varphi_0 \frac{\partial \alpha}{\partial z} = -\frac{u_\alpha}{h_\alpha} + \varphi_0 \frac{\partial \alpha}{\partial z},$$

$$\frac{1}{h_\beta^2} \frac{\partial F}{\partial \beta} = -u \frac{\partial \beta}{\partial r} - w \frac{\partial \beta}{\partial z} + \varphi_0 \frac{\partial \beta}{\partial z} = -\frac{u_\beta}{h_\beta} + \varphi_0 \frac{\partial \beta}{\partial z}.$$

Here,  $h_\alpha$  and  $h_\beta$  are the metric coefficients and  $u_\alpha, u_\beta$  the displacements in  $\alpha, \beta$  coordinates.

Hence the solution finally becomes

$$u_\alpha = -\frac{1}{h_\alpha} \frac{\partial F}{\partial \alpha} + h_\alpha \frac{\partial \alpha}{\partial z} \varphi_0$$

$$u_\beta = -\frac{1}{h_\beta} \frac{\partial F}{\partial \beta} + h_\beta \frac{\partial \beta}{\partial z} \varphi_0,$$

where, as before,

$$F = \frac{\lambda + \mu}{2(\lambda + 2\mu)} z \varphi_0 + \varphi_1, \quad \nabla^2 \varphi_0 = 0, \quad \nabla^2 \varphi_1 = 0.$$

## APPENDIX II

*Solution for a Narrow Rectangular Cross-Section*

When the cross-section of the tore is narrow rectangular, the stresses  $\widehat{zz}$  and  $\widehat{zr}$  may be assumed to vanish and the others to be independent of  $z$ ; the problem is then one of plane stress.

From equations (33) and (36) of Section IV  $\widehat{zz} = 0$  becomes

$$\begin{aligned} -\mu \frac{\partial \varphi_0}{\partial z} + \mu \frac{\lambda + \mu}{\lambda + 2\mu} z \frac{\partial^2 \varphi_0}{\partial z^2} + 2\mu \frac{\partial^2 \varphi_1}{\partial z^2} \\ = \frac{2\lambda\mu K}{\lambda + 2\mu} \log r + \frac{\lambda\mu K}{\lambda + 2\mu} + \lambda K \quad \dots\dots\dots(1) \end{aligned}$$

and  $\widehat{rz} = 0$  becomes

$$\frac{\mu^2}{\lambda + 2\mu} \frac{\partial \varphi_0}{\partial r} - \mu \frac{\lambda + \mu}{\lambda + 2\mu} z \frac{\partial^2 \varphi_0}{\partial r \partial z} - 2\mu \frac{\partial^2 \varphi_1}{\partial r \partial z} = 0 \quad \dots\dots(2)$$

Differentiating (1) with respect to  $r$ , (2) with respect to  $z$  and adding, gives

$$\frac{\partial^2 \varphi_0}{\partial r \partial z} = -\frac{\lambda K}{2(\lambda + \mu)} \frac{1}{r} \quad \dots\dots\dots(3)$$

Hence

$$\frac{\partial \varphi_0}{\partial r} = -\frac{\lambda K}{2(\lambda + \mu)} \frac{z}{r} + f_1(r), \quad \dots\dots\dots(4)$$

$$\frac{\partial \varphi_0}{\partial z} = -\frac{\lambda K}{2(\lambda + \mu)} \log r + f_2(z); \quad \dots\dots\dots(5)$$

$f_1(r)$ ,  $f_2(z)$ , . . . will denote arbitrary functions of integration the form of which is determined subsequently.

From (4) and (5),  $\nabla^2 \varphi_0 = 0$  becomes

$$\frac{1}{r} f_1(r) + \frac{df_1(r)}{dr} + \frac{df_2(z)}{dz} = 0,$$

so that

$$\frac{1}{r} f_1(r) + \frac{df_1(r)}{dr} = -\frac{df_2(z)}{dz} = K_1 \quad \dots\dots\dots(6)$$

$K_1$ ,  $K_2$ , . . . denote arbitrary constants to be determined by the boundary conditions.

From (6),

$$f_1(r) = \frac{1}{2} K_1 r^2 + K_3/r, \quad \dots\dots\dots(7)$$

$$f_2(z) = -K_1 z^2 + K_2 \quad \dots\dots\dots(8)$$

Substituting (7) and (8) in (3) and (4), and using the result in (2) gives

$$\begin{aligned} 2 \frac{\partial^2 \varphi_1}{\partial r \partial z} &= \frac{\mu}{\lambda + 2\mu} \frac{\partial \varphi_0}{\partial r} - \frac{\lambda + \mu}{\lambda + 2\mu} z \frac{\partial^2 \varphi_0}{\partial r \partial z}, \\ &= \frac{\mu}{\lambda + 2\mu} \left( -\frac{\lambda K}{2(\lambda + \mu)} \frac{z}{r} + \frac{1}{2} K_1 r + \frac{K_3}{r} \right) + \frac{\lambda + \mu}{\lambda + 2\mu} \cdot \frac{\lambda K}{2(\lambda + \mu)} \frac{z}{r}, \\ &= \frac{\lambda^2 K}{2(\lambda + \mu)(\lambda + 2\mu)} \frac{z}{r} + \frac{\mu}{\lambda + 2\mu} \left( \frac{1}{2} K_1 r + \frac{K_3}{r} \right). \end{aligned}$$



Hence

$$2 \frac{\partial \varphi_1}{\partial r} = \frac{\lambda^2 K}{4(\lambda + \mu)(\lambda + 2\mu)} \frac{z^2}{r} + \frac{\mu}{\lambda + 2\mu} \left( \frac{1}{2} K_1 r z + K_3 \frac{z}{r} \right) + f_3(r), \quad \dots (9)$$

$$2 \frac{\partial \varphi_1}{\partial z} = \frac{\lambda^2 K}{2(\lambda + \mu)(\lambda + 2\mu)} z \log r + \frac{\mu}{\lambda + 2\mu} (\frac{1}{2} K_1 + K_3 \log r) + f_4(z) \dots (10)$$

From (9) and (10),  $\nabla^2 \varphi_1 = 0$

becomes

$$\frac{\lambda^2 K}{2(\lambda + \mu)(\lambda + 2\mu)} \log r + \frac{f_3(r)}{r} + \frac{df_3(r)}{dr} = - \frac{\mu K_1}{\lambda + 2\mu} z - \frac{df_4(z)}{dz} = K_4,$$

therefore

$$\frac{d}{dr} [r f_3(r)] = r K_4 - \frac{\lambda^2 K}{2(\lambda + \mu)(\lambda + 2\mu)} r \log r,$$

therefore

$$f_3(r) = \frac{1}{2} r K_4 - \frac{\lambda^2 K}{2(\lambda + \mu)(\lambda + 2\mu)} \cdot \frac{1}{2} r (\log r - \frac{1}{2}) + \frac{K_5}{r}, \quad \dots (11)$$

and

$$f_4(z) = - z K_4 - \frac{\mu K_1}{2(\lambda + 2\mu)} z^2 + K_6 \quad \dots (12)$$

Substituting (11) of Section IV and (12) in (9) and (10) and the results in (33) and (36) gives the expressions for the stresses; since the stress-system is independent of  $z$ , terms containing  $z$  will be omitted. The stresses  $\widehat{rr}$  and  $\widehat{\delta\delta}$  thus become

$$\begin{aligned} \widehat{rr} = & \left( \frac{2\mu K(\lambda + \mu)}{\lambda + 2\mu} - \frac{\lambda^2 \mu K}{4(\lambda + 2\mu)(\lambda + \mu)} \right) \log r + \frac{\mu K_5}{r^2} + \frac{\lambda^2 \mu K}{8(\lambda + 2\mu)(\lambda + \mu)} \\ & + \frac{\lambda \mu K_2}{\lambda + 2\mu} - \frac{1}{2} K_4 \mu + K(\lambda + \mu), \quad \dots (13) \end{aligned}$$

$$\begin{aligned} \widehat{\delta\delta} = & \left( \frac{2\mu K(\lambda + \mu)}{\lambda + 2\mu} - \frac{\lambda^2 \mu K}{4(\lambda + 2\mu)(\lambda + \mu)} \right) \log r - \frac{\mu K_5}{r^2} - \frac{\lambda^2 \mu K}{8(\lambda + 2\mu)(\lambda + \mu)} \\ & + \frac{\lambda \mu K_2}{\lambda + 2\mu} - \frac{1}{2} K_4 \mu + \frac{\lambda \mu K}{\lambda + 2\mu} + K(\lambda + 2\mu) \quad \dots (14) \end{aligned}$$

The constants in (13) and (14) have now to be determined by the boundary conditions

$$\widehat{rr} = 0 \text{ for } r = A \text{ and } r = B. \quad \dots (15)$$

It is easily shown that (15) gives for the constants

$$K_5 = \frac{A^2 B^2}{\mu(A^2 - B^2)} \left( \frac{2\mu K(\lambda + \mu)}{\lambda + 2\mu} - \frac{\lambda^2 \mu K}{4(\lambda + 2\mu)(\lambda + \mu)} \right) \log \frac{A}{B}, \quad \dots (16)$$

and

$$\begin{aligned} \frac{1}{2} K_4 \mu - \frac{2\mu K_2}{\lambda + 2\mu} = & \left( \frac{2\mu K(\lambda + \mu)}{\lambda + 2\mu} - \frac{\lambda^2 \mu K}{4(\lambda + 2\mu)(\lambda + \mu)} \right) \frac{A^2 \log A - B^2 \log B}{A^2 - B^2} \\ & + \frac{\lambda^2 \mu K}{8(\lambda + 2\mu)(\lambda + \mu)} + K(\lambda + \mu). \quad \dots (17) \end{aligned}$$

It is convenient at this stage to introduce the elastic constants  $\sigma$  and  $E$  by the equations

$$\lambda = \frac{E\sigma}{(1+\sigma)(1-2\sigma)}, \quad \mu = \frac{E}{2(1+\sigma)}.$$

With the constants (16) and (17), the expressions for the stresses (13) and (14) finally reduce to

$$\widehat{rr} = -\frac{E(2-\sigma^2)}{4(1-\sigma^2)} \frac{K}{A^2-B^2} \left( \frac{A^2 B^2}{r^2} \log \frac{B}{A} + \frac{A^2 \log A}{r} + B^2 \log \frac{r}{B} \right) \quad (18)$$

$$\widehat{\delta\delta} = \frac{E(2-\sigma^2)}{r(1-\sigma^2)} \frac{K}{A^2-B^2} \left[ \frac{A^2 B^2}{r^2} \log \frac{B}{A} - B^2 \log \frac{r}{B} - A^2 \log \frac{A}{r} - (B^2 - A^2) \right] \quad \dots\dots\dots (19)$$

The value of  $K$  is proportional to the applied flexural couple and is found by substituting  $\widehat{\delta\delta}$  from (19) in equation (43) of Section VI and integrating.

The solutions (18), (19) for the narrow rectangular cross-section agree with that of Golovin (quoted by Timoshenko<sup>10</sup>) and that of Southwell<sup>(2)</sup>.

# ON THE SOLUTION OF THE EQUILIBRIUM EQUATIONS OF ELASTICITY IN GENERAL CURVILINEAR COORDINATES

By W. FREIBERGER\*

[*Manuscript received July 22, 1949*]

## *Summary*

A system of stress functions consisting of a harmonic vector and a harmonic scalar potential is introduced into the equations of equilibrium of an isotropic elastic body in arbitrary curvilinear coordinates and tensor form. The components of displacement and of the stress tensor are expressed in terms of these potentials and it is shown that the total number of independent components of the potentials is reducible to three.

The tensor equations are then specialized to orthogonal curvilinear coordinates and expressions derived for the "physical" components of displacement and stress in terms of the stress functions.

Reference is made to the relation between the present approach and previous ones.

## I. INTRODUCTION

Direct solutions of the equations of equilibrium of an elastic body in the general case present great difficulties. Recourse is therefore had to the method of stress-functions which was first introduced by Airy for the two-dimensional case.

Amongst approaches to the three-dimensional problem one may distinguish between those starting from Cauchy's equations in terms of stresses and others which take Navier's displacement equations as their point of departure.

Maxwell was the first to formulate stress-functions for the solution of Cauchy's equations; another set is due to Morera. Galerkin(1), Papkovitch(2), and Neuber(3) treated the displacement equations in Cartesian coordinates by this approach.

In this paper a system of stress functions is introduced into Navier's equilibrium equations in tensor form for general curvilinear coordinates. It consists of a harmonic scalar and a harmonic vector potential-function and it is shown that one of these four components can always be eliminated. The potentials must be made to satisfy the boundary conditions on the surface of the elastic body; these conditions are in general simultaneous in the three stress-functions so that one does not get a reduction to Dirichlet or Neumann problems.

The curvilinear coordinate system is next restricted to be orthogonal and expressions for the "physical" components of displacement and stress deduced from the tensor equations.

Finally, the present approach is related to the equations of Galerkin and Papkovitch-Neuber.

\*Aeronautical Research Laboratories, Department of Supply and Development, Melbourne.

## II. SOLUTION IN TENSOR FORM

Navier's equations of equilibrium of an isotropic elastic body are

$$(\lambda + \mu) \frac{\partial \theta}{\partial q^r} + \mu g^{st} u_{r,st} = 0, \quad \dots\dots\dots(1)$$

where  $\lambda, \mu$  are Lamé's elastic constants of the material,

$u_r$  is the covariant component of displacement,

$q^r$  is the contravariant component of the coordinates,

$g^{st}$  is the contravariant component of the fundamental (metric) tensor,

$\theta$  is the cubical dilatation :

$$\theta = u^j_{,j} \quad \dots\dots\dots(2)$$

The Einstein summation convention will be employed throughout this section ; that is, summation is implied over repeated suffices. Covariant differentiation is denoted by a comma followed by the suffices with respect to which the differentiation is carried out.

The function  $F$  is now introduced by the substitution

$$(\lambda + \mu)\theta = g^{st}F_{,st} \quad \dots\dots\dots(3)$$

Differentiating this covariantly with respect to  $q^r$

$$(\lambda + \mu) \frac{\partial \theta}{\partial q^r} = g^{st} \left( \frac{\partial F}{\partial q^r} \right)_{,st} \quad \dots\dots\dots(4)$$

since covariant differentiation of a scalar is ordinary differentiation, and

$$g^{st}_{,r} = 0.$$

Also, covariant differentiation is here commutative since it is assumed that the space used admits of a Cartesian coordinate system : this implies the vanishing of the Riemann-Christoffel curvature tensor (i.e. "flat" space)—the condition for commutative covariant differentiation.

Substituting (4) in (1)

$$g^{st} \left( \frac{\partial F}{\partial q^r} + \mu u_r \right)_{,st} = 0$$

so that

$$\mu u_r = \varphi_r - \frac{\partial F}{\partial q^r}, \quad \dots\dots\dots(5)$$

where  $\varphi_r$  is a vector satisfying Laplace's equation :

$$g^{st} \varphi_{r,st} = 0 \quad \dots\dots\dots(6)$$

From (5) the contravariant components of displacement may be obtained :

$$\mu u^j = \mu g^{ij} u_r = -g^{ij} \frac{\partial F}{\partial q^i} + \varphi^j \quad \dots\dots\dots(7)$$

Substituting this in (2)

$$\mu \theta = \left( -g^{ij} \frac{\partial F}{\partial q^i} + \varphi^j \right)_{,j} \quad \dots\dots\dots(8)$$



Substituting (8) in (3)

$$\frac{\lambda + \mu}{\mu} \left( -g^{ij} \frac{\partial F}{\partial q^i} + \varphi^j \right)_{,j} = g^{ij} F_{,ij} = \left( g^{ij} \frac{\partial F}{\partial q^i} \right)_{,j},$$

therefore

$$\left( \frac{\lambda + \mu}{\mu} + 1 \right) g^{ij} F_{,ij} - \frac{\lambda + \mu}{\mu} \varphi^j_{,j} = 0,$$

that is

$$\left[ \left( g^{ij} \frac{\partial F}{\partial q^i} - \frac{\lambda + \mu}{\lambda + 2\mu} \varphi^j \right) \right]_{,j} = 0,$$

that is

$$g^{ij} F_{,ij} = \frac{\lambda + \mu}{\lambda + 2\mu} \varphi^j_{,j} \quad \dots\dots\dots (9)$$

A particular solution of (9) is

$$F = \frac{\lambda + \mu}{\lambda + 2\mu} q^r \varphi_r; \quad \dots\dots\dots (10)$$

this may be verified as follows :

From (10) :

$$F_{,i} = \frac{\lambda + \mu}{\lambda + 2\mu} (\varphi_i + q^r \varphi_{r,i}),$$

therefore

$$\begin{aligned} g^{ij} F_{,ij} &= \frac{\lambda + \mu}{\lambda + 2\mu} (\varphi^j_{,j} + q^r g^{ij} \varphi_{r,ij}), \\ &= \frac{\lambda + \mu}{\lambda + 2\mu} \varphi^j_{,j} \end{aligned}$$

from equation (6), as asserted in equation (9).

To the particular solution (10) of (9) has to be added the general solution of

$$g^{ij} F_{,ij} = 0,$$

which will be denoted by  $\varphi_0$ , so as to give the general solution of (9). The latter therefore has the form

$$F = \varphi_0 + \frac{\lambda + \mu}{\lambda + 2\mu} q^r \varphi_r \quad \dots\dots\dots (11)$$

Conversely, we shall now show that, when the  $u_r$  are specified by (5) with  $F$  determined by (11), with  $\varphi_0$  and the  $\varphi_r$  harmonic, then equations (1) and (3) are satisfied.

We first show that

$$F = \varphi_0 + \frac{\lambda + \mu}{\lambda + 2\mu} q^r \varphi_r \quad \dots\dots\dots (11)$$

satisfies

$$(\lambda + \mu) \theta = g^{st} F_{,st} \quad \dots\dots\dots (3)$$

From (11)

$$g^{st}F_{,st} = \frac{\lambda + \mu}{\lambda + 2\mu} \varphi^j_{,j}$$

so that the following equation has to be verified :

$$\begin{aligned} (\lambda + 2\mu)\theta &= \varphi^j_{,j}, \\ \text{that is } (\lambda + 2\mu)u^j_{,j} &= \varphi^j_{,j}, \end{aligned}$$

now from (5)

$$\varphi^j_{,j} = \mu u^j_{,j} + F^j_{,j}; \quad \dots\dots\dots(5')$$

but

$$\begin{aligned} F^j_{,j} &= \frac{\lambda + \mu}{\lambda + 2\mu} (\varphi^j_{,j} + q^r \varphi_{r,j}), \\ &= \frac{\lambda + \mu}{\lambda + 2\mu} \varphi^j_{,j} \quad \text{since } \varphi_{r,j} = 0; \end{aligned}$$

hence from (5')

$$\varphi^j_{,j} = (\lambda + 2\mu)u^j_{,j},$$

which verifies equation (3). Now to verify

$$(\lambda + \mu) \frac{\partial \theta}{\partial q^r} + \mu g^{st} u_{r,st} = 0, \quad \dots\dots\dots(1)$$

substitute in (1) for  $\theta$  from (3) and for  $u_r$  from (5), to obtain for the left-hand side of (1)

$$\frac{\partial}{\partial q^r} (g^{st} F_{,st}) + g^{st} \varphi_{r,st} - g^{st} \left( \frac{\partial F}{\partial q^r} \right)_{,st} = g^{st} F_{,str} - g^{st} F_{,str} = 0,$$

from (6), from the commutative property of covariant differentiation, and because  $g^{st}{}_{,r} = 0$ ; this verifies equation (1).

It has now been established that any solution of (1) for  $u_r$  is expressible in the form (5), with  $F$  specified in the form (11); i.e. that

$$\mu u_r = \varphi_r - \frac{\partial F}{\partial q^r}, \quad (r = 1, 2, 3)$$

with

$$F = \varphi_0 + \frac{\lambda + \mu}{\lambda + 2\mu} q^r \varphi_r;$$

the scalar  $\varphi_0$  and the vector  $(\varphi_1, \varphi_2, \varphi_3)$  are harmonic.

To show that one of the functions,  $\varphi_1, \varphi_2, \varphi_3$  can always be eliminated from the solution, we introduce the new functions,  $\varphi'_0, \varphi'_1, \varphi'_2, \varphi'_3$  by the equations :

$$\varphi_0 = \varphi'_0 + \varphi'_3 - \frac{\lambda + \mu}{\lambda + 2\mu} q^r \frac{\partial \varphi'_3}{\partial q^r},$$

$$\varphi_1 = \varphi'_1 + \frac{\partial \varphi'_3}{\partial q^1},$$

$$\varphi_2 = \varphi'_2 + \frac{\partial \varphi'_3}{\partial q^2},$$

$$\varphi_3 = \frac{\partial \varphi'_3}{\partial q^3}.$$

First we see that scalar function  $\varphi_0$  and the vector function  $(\varphi_1', \varphi_2', \varphi_3')$  defined by these equations can again be chosen harmonic, if  $\varphi_0$  and  $(\varphi_1, \varphi_2, \varphi_3)$  are harmonic ; in fact,

$$g^{st}\varphi_{1,st} = g^{st}\varphi_{1,st}' + g^{st}\left(\frac{\partial\varphi_3'}{\partial q_1}\right)_{,st} = 0, \text{ since } g^{st}\varphi_{r,st} = 0 \quad (r = 1, 2, 3).$$

This is satisfied if

$$g^{st}\varphi_{1,st}' = 0$$

since

$$g^{st}\left(\frac{\partial\varphi_3'}{\partial q^2}\right)_{,st} = (g^{st}\varphi_{3,st}')_{,1}.$$

Similarly,  $g^{st}\varphi_{2,st}' = g^{st}\varphi_{3,st}' = 0$  ;

hence  $g^{st}\varphi_{r,st}' = 0$  and so  $(\varphi_1', \varphi_2', \varphi_3')$  is harmonic. Now for  $\varphi_0'$ :

$$g^{st}\varphi_{0,st} = g^{st}\varphi_{0,st}' + g^{st}\varphi_{3,st}' - \frac{\lambda + \mu}{\lambda + 2\mu} g^{st}(q^r\varphi_{3,r}')_{,st} = 0.$$

The first two terms vanish if  $\varphi_0'$  and  $(\varphi_1', \varphi_2', \varphi_3')$  are harmonic ; the last becomes

$$\begin{aligned} g^{st}(\varphi_{3,s} + q^r\varphi_{3,rs})_{,t} &= g^{st}(\varphi_{3,st} + \varphi_{3,st} + q^r\varphi_{3,rst}) \\ &= 2g^{st}\varphi_{3,st} + q^r(g^{st}\varphi_{3,st})_{,r} = 0 \end{aligned}$$

since  $g^{st}\varphi_{r,st} = 0$ .

This proves that  $\varphi_0'$  and  $(\varphi_1', \varphi_2', \varphi_3')$  are again harmonic.

We further introduce the function  $F'$  by the equation

$$F = F' + \varphi_3'.$$

Then clearly

$$\begin{aligned} F' &= F - \varphi_3' = \varphi_0 + \frac{\lambda + \mu}{\lambda + 2\mu} q^r\varphi_r - \varphi_3', \\ &= \varphi_0' + \frac{\lambda + \mu}{\lambda + 2\mu} (q^1\varphi_1' + q^2\varphi_2'), \end{aligned}$$

and the displacements become

$$\begin{aligned} \mu u_1 &= \varphi_1 - \frac{\partial F}{\partial q^1} = \varphi_1' - \frac{\partial F'}{\partial q^1}, \\ \mu u_2 &= \varphi_2 - \frac{\partial F}{\partial q^2} = \varphi_2' - \frac{\partial F'}{\partial q^2}, \\ \mu u_3 &= \varphi_3 - \frac{\partial F}{\partial q^3} = - \frac{\partial F'}{\partial q^3}, \end{aligned}$$

which do not contain the function  $\varphi_3'$ . Omitting the dashes it may therefore be said that the displacements are expressible in terms of three functions  $\varphi_0, \varphi_1, \varphi_2$  thus

$$\mu u_1 = \varphi_1 - \frac{\partial F}{\partial q^1}, \mu u_2 = \varphi_2 - \frac{\partial F}{\partial q^2}, \mu u_3 = - \frac{\partial F}{\partial q^3},$$

where

$$F = \varphi_0 + \frac{\lambda + \mu}{\lambda + 2\mu} (q^1 \varphi_1 + q^2 \varphi_2);$$

the scalar function  $\varphi_0$  and the vector function  $(\varphi_1, \varphi_2)$  are both harmonic.

In the following pages, the four functions  $\varphi_0, \varphi_1, \varphi_2, \varphi_3$  will be retained, because it depends on the particular problem which of  $\varphi_1, \varphi_2, \varphi_3$  can most conveniently be equated to zero.

The equilibrium stresses will now be found in terms of the stress potentials.

The components of the stress-tensor may be written in terms of the displacements as follows :

$$\tau_{rs} = \lambda g_{rs} u^j{}_{,j} + \mu (u_{r,s} + u_{s,r}), \quad \dots\dots\dots(12)$$

which is the generalized Hooke's law.

From equation (7)

$$\begin{aligned} \mu u^j{}_{,j} &= -g^{ij} F_{,ij} + \varphi^j{}_{,j} \\ \mu u_{r,s} &= -F_{,rs} + \varphi_{r,s} \\ \mu u_{s,r} &= -F_{,sr} + \varphi_{s,r} \end{aligned}$$

so that in equation (12)

$$\tau_{rs} = \frac{\lambda}{\mu} g_{rs} (-g^{ij} F_{,ij} + \varphi^j{}_{,j}) + (-2F_{,rs} + \varphi_{r,s} + \varphi_{s,r}) \quad \dots\dots(13)$$

since  $F_{,rs} = F_{,sr}$ .

But

$$\varphi^j{}_{,j} = \frac{\lambda + 2\mu}{\lambda + \mu} g^{ij} F_{,ij} \quad \text{from (9),}$$

so that (13) becomes

$$\tau_{rs} = \frac{\lambda}{\lambda + \mu} g_{rs} g^{ij} F_{,ij} - 2F_{,rs} + \varphi_{r,s} + \varphi_{s,r} \quad \dots\dots\dots(14)$$

which is the required solution.

It will be convenient to write equation (14) in terms of ordinary differentiations. To this end it must be recalled that for any covariant vector  $X_r$  the first covariant derivative is

$$X_{r,s} = \frac{\partial X_r}{\partial q^s} - \left\{ \begin{matrix} m \\ rs \end{matrix} \right\} X_m,$$

where  $\left\{ \begin{matrix} m \\ rs \end{matrix} \right\}$  is the Christoffel symbol of the second kind defined by the formula

$$\left\{ \begin{matrix} m \\ rs \end{matrix} \right\} = \frac{1}{2} g^{mp} \left( \frac{\partial g_{rp}}{\partial q^s} + \frac{\partial g_{ps}}{\partial q^r} - \frac{\partial g_{rs}}{\partial q^p} \right).$$

The second covariant derivative of a covariant vector is

$$\begin{aligned} X_{r, st} &= \frac{\partial^2 X_r}{\partial q^s \partial q^t} - \left\{ \begin{matrix} m \\ rs \end{matrix} \right\} \frac{\partial X_m}{\partial q^t} - \left\{ \begin{matrix} m \\ rt \end{matrix} \right\} \frac{\partial X_m}{\partial q^s} - \left\{ \begin{matrix} m \\ st \end{matrix} \right\} \frac{\partial X_r}{\partial q^m} \\ &\quad - X_m \left[ \frac{\partial}{\partial q^t} \left\{ \begin{matrix} m \\ rs \end{matrix} \right\} + \left\{ \begin{matrix} m \\ rp \end{matrix} \right\} \left\{ \begin{matrix} p \\ st \end{matrix} \right\} - \left\{ \begin{matrix} m \\ sp \end{matrix} \right\} \left\{ \begin{matrix} p \\ rt \end{matrix} \right\} \right]. \end{aligned}$$



Since

$$g^{ij}F_{,ij} \equiv \nabla^2 F,$$

the components (14) of the stress tensor finally become

$$\left. \begin{aligned} \tau_{rs} &= \frac{\lambda}{\lambda + \mu} g_{rs} \nabla^2 F - 2 \frac{\partial^2 F}{\partial q^r \partial q^s} + \frac{\partial \varphi_r}{\partial q^s} + \frac{\partial \varphi_s}{\partial q^r} + 2 \left( \frac{\partial F}{\partial q^m} - \varphi_m \right) \left\{ \begin{matrix} m \\ rs \end{matrix} \right\}, \\ F &= \varphi_0 + \frac{\lambda + \mu}{\lambda + 2\mu} q^m \varphi_m, \quad \nabla^2 \varphi_0 = 0, \quad g^{ij} \varphi_{r,ij} = 0. \end{aligned} \right\} \quad (15)$$

### III. PHYSICAL COMPONENTS OF STRESS

In classical elasticity theory stress components are referred to Cartesian axes coinciding in direction with the curvilinear system at the point considered and become particularly simple for orthogonal coordinates. These components are called the "physical" components of stress and they will now be obtained from the tensor components given by equation (15).

In this section  $q^r$  and  $q_r$  will denote orthogonal curvilinear coordinates and the summation convention will be dropped.

Then

$$g_{mn} = g^{mn} = 0 \quad (m \neq n), \quad g_{mm} = \frac{1}{g^{mm}} = h_m^2,$$

where the line-element is given by

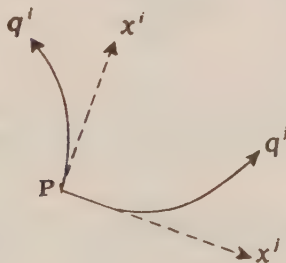
$$ds^2 = h_1^2 (dq^1)^2 + h_2^2 (dq^2)^2 + h_3^2 (dq^3)^2.$$

For a Cartesian coordinate system  $x^r$  at the point considered with the directions of its axes the same as the directions of the coordinate lines  $q^r$  at the point

$$\begin{aligned} ds^2 &= (dx^1)^2 + (dx^2)^2 + (dx^3)^2, \\ &= \left( \frac{dx^1}{dq^1} \right)^2 (dq^1)^2 + \left( \frac{dx^2}{dq^2} \right)^2 (dq^2)^2 + \left( \frac{dx^3}{dq^3} \right)^2 (dq^3)^2, \end{aligned}$$

since

$$\partial x^i / \partial q^j = 0 \quad \text{if } i \neq j.$$



Clearly, therefore,

$$\frac{\partial x^1}{\partial q^1} = h_1, \quad \frac{\partial x^2}{\partial q^2} = h_2, \quad \frac{\partial x^3}{\partial q^3} = h_3.$$

By the transformation law for 2nd order covariant tensor components,

$$\begin{aligned}\tau_{rs} &= \sum_{\lambda, \nu} \frac{\partial x^\lambda}{\partial q^r} \frac{\partial x^\nu}{\partial q^s} \widehat{x_\lambda x_\nu} \\ &= \frac{\partial x^r}{\partial q^r} \frac{\partial x^s}{\partial q^s} \widehat{x_r x_s} \quad \text{since } \partial x^i / \partial q^j = 0 \text{ if } i \neq j \\ &= h_r h_s \widehat{x_r x_s},\end{aligned}$$

where  $\widehat{x_r x_s}$  are the physical components of stress.

Similarly, for covariant vectors

$$\begin{aligned}\varphi_r &= \sum_{\lambda} \frac{\partial x^\lambda}{\partial q^r} \bar{\varphi}_\lambda \\ &= \frac{\partial x^r}{\partial q^r} \bar{\varphi}_r \\ &= h_r \bar{\varphi}_r,\end{aligned}$$

where  $\bar{\varphi}_r$  are physical components.

Substituting these results into equation (15) gives for the solution

$$\begin{aligned}h_r h_s \widehat{x_r x_s} &= \frac{\lambda}{\lambda + \mu} \delta_{rs} h_r h_s \nabla^2 F - 2 \frac{\partial^2 F}{\partial q^r \partial q^s} \\ &\quad + \frac{\partial}{\partial q^s} (h_r \bar{\varphi}_r) + \frac{\partial}{\partial q^r} (h_s \bar{\varphi}_s) + 2 \sum_m \left( \frac{\partial F}{\partial q^m} - h_m \bar{\varphi}_m \right) \left\{ \begin{matrix} m \\ rs \end{matrix} \right\},\end{aligned}$$

where

$$F = \varphi_0 + \frac{\lambda + \mu}{\lambda + 2\mu} \sum_m q^m h_m \bar{\varphi}_m.$$

Hence, finally, omitting the bars on  $\varphi_r$

$$\begin{aligned}\widehat{x_r x_s} &= \frac{\lambda}{\lambda + \mu} \delta_{rs} \nabla^2 F - \frac{2}{h_r h_s} \frac{\partial^2 F}{\partial q^r \partial q^s} + \frac{1}{h_s} \frac{\partial \varphi_r}{\partial q^s} + \frac{1}{h_r} \frac{\partial \varphi_s}{\partial q^r} \\ &\quad + \frac{1}{h_r h_s} \left[ q_r \frac{\partial h_r}{\partial q^r} + q_s \frac{\partial h_s}{\partial q^r} + 2 \sum_m \left( \frac{\partial F}{\partial q^m} - h_m \varphi_m \right) \left\{ \begin{matrix} m \\ rs \end{matrix} \right\} \right] \quad \dots\dots(16)\end{aligned}$$

where

$$F = \varphi_0 + \frac{\lambda + \mu}{\lambda + 2\mu} \sum_{m=1}^3 h_m q^m \varphi_m \quad \dots\dots\dots(17)$$

Transformation into physical components gives for the displacements, equation (5),

$$\mu u_r = \varphi_r - \frac{1}{h_r} \frac{\partial F}{\partial q^r} \quad \dots\dots\dots(18)$$

The scalar function  $\varphi_0$  and the vector function  $(\varphi_1, \varphi_2, \varphi_3)$  satisfy the equations

$$\nabla^2 \varphi_0 = 0, \quad \nabla^2 (h_1 \varphi_1, h_2 \varphi_2, h_3 \varphi_3) = 0 \quad \dots\dots\dots(19)$$

## IV. RELATION TO PREVIOUS SOLUTIONS

By adapting the notation used in previous solutions to the tensor notation here used, these solutions can be related to the present one.

(i) *Galerkin*

Galerkin's solution can be written

$$\mu u_r = \frac{\lambda + 2\mu}{\lambda + \mu} g^{ij} \Phi_{r,ij} - \Phi^i_{,ir},$$

where  $\Phi_r$  is the biharmonic "Galerkin vector."

Comparing this with equation (5)

$$\mu u_r = \varphi_r - F_{,r} \dots\dots\dots (5)$$

it is seen that the Galerkin vector is related to the present solution functions by the equations

$$\begin{aligned} \varphi_r &= \frac{\lambda + 2\mu}{\lambda + \mu} g^{ij} \Phi_{r,ij} \\ F_{,r} &= \Phi^i_{,ir}, \text{ therefore } F = \Phi^i_{,i}. \end{aligned}$$

In vector notation these relations become

$$\begin{aligned} \vec{\varphi} &= \frac{\lambda + 2\mu}{\lambda + \mu} \nabla^2 \vec{\Phi}, \\ F &= \text{div } \vec{\Phi}. \end{aligned}$$

(ii) *Papkovitch-Neuber*:

In the present notation the Papkovitch-Neuber solution becomes

$$\mu u_r = \frac{\lambda + 2\mu}{\lambda + \mu} \Psi_r - \frac{1}{2} P_{,r};$$

comparing this with equation (5) gives the relations

$$\varphi_r = \frac{\lambda + 2\mu}{\lambda + \mu} \Psi_r, \quad F = \frac{1}{2} P.$$

Neuber's stress-functions are related to Cartesian coordinates; his expressions(4) for the stresses can therefore be obtained from the present solution if the functions  $\varphi_0, \varphi_x, \varphi_y, \varphi_z$  are transformed into Cartesians to become  $\varphi_0, \varphi_x, \varphi_y, \varphi_z$ . The transformation equations are

$$\begin{aligned} \varphi_0 &= \varphi_0 \\ h_r \varphi_r &= \varphi_x \frac{\partial x}{\partial q^r} + \varphi_y \frac{\partial y}{\partial q^r} + \varphi_z \frac{\partial z}{\partial q^r} \quad (r = 1, 2, 3). \end{aligned}$$

Substituting these relations into equation (16) gives the equations for stresses in an orthogonal curvilinear coordinate system in terms of the four harmonic functions  $\varphi_0, \varphi_x, \varphi_y, \varphi_z$ , and similarly for the special coordinate systems of Section IV.

## V. CONCLUSION

The solution of the three-dimensional equilibrium equations in terms of stress functions given in Section II for general tensor coordinates was specialized in Section III to orthogonal curvilinears. From this, expressions can be derived for the physical component of stress and displacement in particular coordinate systems; this is done in reference (5) for Cartesian, cylindrical, spherical, ellipsoidal, and toroidal coordinates.

The method of approach presented here is appropriate to three-dimensional elasticity problems in which the body under consideration is bounded by as few coordinate surfaces as possible; it reduces the problem to a number of potential problems for the functions  $\varphi_0$  and  $\varphi_r$  ( $r = 1, 2, 3$ ), which assume prescribed values on the boundaries but enter simultaneously into the boundary conditions. A reduction to a Dirichlet or Neumann-problem will therefore in general not be possible.

It has been shown that one of the functions  $\varphi_r$  ( $r = 1, 2, 3$ ) can always be chosen to vanish identically. In practice, it will often be found that fewer than even the three functions suffice for the solution, depending on the degree of symmetry of the problem. For instance, in the problem of the flexure of incomplete tores(6) there is considerable simplification. Because the coordinates are cylindrical, the equation  $g^{ij}q_{3,ij} = 0$  reduces to the ordinary scalar  $\nabla^2 q_3 = 0$ . It is also possible to put  $q_2 = 0$  and to take  $\varphi_1$ ,  $\varphi_3$  independent of the angular coordinate  $\theta$ . The solution can then be expressed in terms of  $q_0$ ,  $q_3$ , both potential functions.

## VI. ACKNOWLEDGMENTS

The work reported here was carried out as part of the research programme of the Aeronautical Research Laboratories. The author is indebted to the Chief Scientist, Department of Supply and Development, for permission to publish this paper.

## VII. REFERENCES

- (1) GALERKIN, B.—*C.R. Acad. Sci. U.R.S.S.* A **14**: 353 (1930).
- (2) PAFKOVITCH, P. F.—*C.R. Acad. Sci. Paris* **195**: 513, 754 (1932).
- (3) NEUBER, H.—*Z. Angw. Math. Mech.* **14**: 203 (1934).
- (4) NEUBER, H.—"Kerbspannungslehre." (Julius Springer: Berlin, 1937).
- (5) FREIBERGER, W.—Aero. Res. Lab. Rep. S.M. 127, Melbourne, Australia.
- (6) FREIBERGER, W., and SMITH, R. C. T.—*Aust. J. Sci. Res.* A **2**: 469 (1949).



# COSMIC RAY MEASUREMENTS BETWEEN AUSTRALIA AND JAPAN

By P. G. LAW\*, C. D. MCKENZIE\*, and H. D. RATHGEBER\*

[Manuscript received April 29, 1949]

## Summary

Cosmic ray results obtained on a journey to and fro between Australia and Japan are described. The apparatus was that used previously on the H.M.A.S. *Wyatt Earp*(1), but since modified, and extended to record showers. A latitude effect of 20 per cent. was found for rays of vertical incidence. The difference between electron and meson components did not exceed the statistical errors of 2 per cent. For wide angle coincidence telescope measurements the latitude effect reduces to 13 per cent. Further it has been found that a latitude effect of some 10 per cent. exists for extensive showers of an average spread of 1 metre. An observed latitude effect of penetrating extensive showers falls, however, within the standard deviation of the measurements. The above results were not corrected for barometric and temperature effects nor for variations in height of the meson-producing layer. Evidence is presented of the existence of a semi-diurnal variation at the equator and at Kure of opposite phase to the barometric pressure variations and of about twice the coefficient of the normal barometric effect. The maximum of the diurnal period occurs at Kure at 2 hours local time and at the equator at 19 hours. A temporary increase of between 2 and 3 per cent. in cosmic ray intensity during the period July 29-31, 1948, is noted.

## I. INTRODUCTION

During the cruise of H.M.A.S. *Wyatt Earp* early in 1948 in Antarctic waters, cosmic ray records covering the latitude range  $67-38^{\circ}$  S. were obtained(1). When the ship returned to Melbourne it was decided to extend the measurements to cover tropical latitudes by sending the equipment on a return trip to Japan aboard M.V. *Duntroon*.

The apparatus was installed with its top 18 in. below a sports deck consisting of steel plate 1 cm. thick and wood 4 cm. thick. The *Duntroon* sailed from Sydney on July 2, 1948, called at Dreger Harbour, New Guinea, on July 8, and reached Kure, Japan ( $34^{\circ}$  N.  $132^{\circ}$  E.) on July 18. Leaving Kure on July 23, the ship again called at Dreger Harbour and then headed for Auckland, New Zealand, where she arrived on August 9. Six days later the *Duntroon* sailed for Australia and berthed in Sydney Harbour on August 20 (Fig. 1).

\*Physics Department, University of Melbourne. Two of the authors (P.G.L. and C.D.McK.) accompanied the ship and operated the equipment.

## II. DESCRIPTION OF EQUIPMENT

The *Wyatt Earp* equipment was modified before being placed on the *Duntroon* as a result of experience gained during the first cruise. The scaling circuits were converted from pentode to triode type. The addition of further scaling stages which increased the scaling ratios to 8, 4, and 32 for counts 1, 2, and 3 respectively, prevented the recurrence of earlier troubles with telephone registers. Considering that the standard errors ( $\sqrt{N}$ ) of the hourly counting rates amount to 45, 36, and 105 respectively, the errors introduced by the increased scaling ratios can still be neglected.

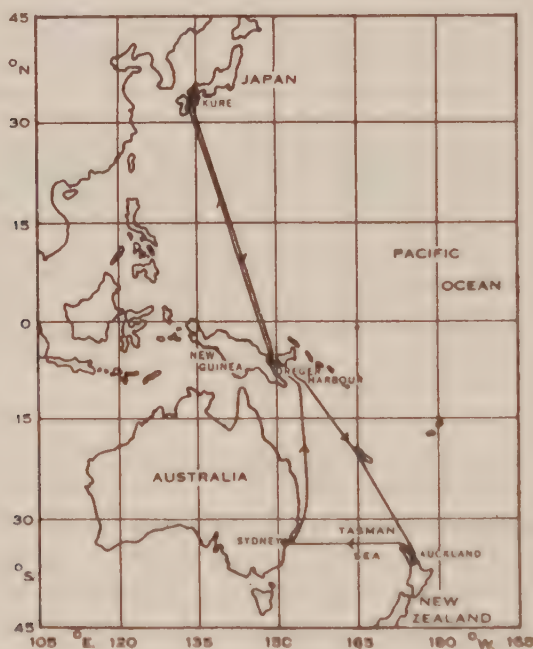


Fig. 1.—Track chart of M.V. *Duntroon*.

New circuits were included in the apparatus to record extensive showers over a base of 1 m. (Fig. 2). Thus, as well as counts 1, 2, and 3 which were the same as in the *Wyatt Earp* apparatus, there were, in addition, counts 4 and 5. Count 4 registered all extensive air showers (fivefold coincidences of trays 1, 2, 3, 5, and 6), while count 5 (coincidences between trays 2, 3, 4, 5, and 6) registered extensive air showers of which at least one particle had penetrated 10 cm. of lead, i.e. penetrating showers.

As some of the counters previously used had deteriorated and as no time was

available before departure to prepare sufficient new ones, the number of counters in each tray was reduced from 6 to 4.

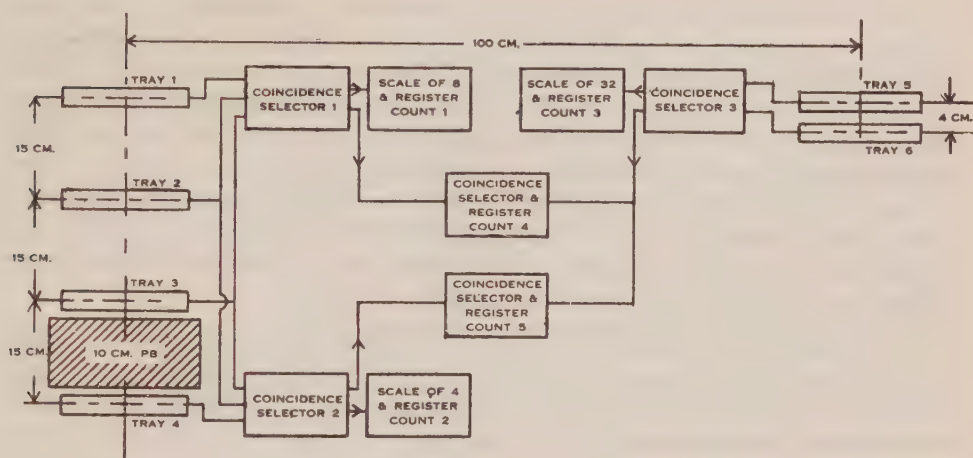


Fig. 2.—Diagram of equipment. Each tray contains four Geiger counters of 18 cm. effective length and 3.1 to 3.3 cm. diameter.

### III. RESULTS AND DISCUSSION

Since its discovery by Clay, the latitude effect has been the subject of numerous investigations(2, 3). Part of the observed variation with latitude of cosmic ray intensity at sea-level is attributed to the deflection of charged primary cosmic ray particles by the earth's magnetic field, and this is referred to as the "geomagnetic latitude effect." Superimposed upon this effect are variations with latitude due to processes which occur during the passage of the cosmic ray particles through the atmosphere.

Present theory assumes that mesons are produced by primary cosmic rays in the upper atmosphere maximum production occurring at a level of constant atmospheric pressure, equal to approximately 100 mb. at a height of about 16 km. above sea-level. This "meson-producing" layer varies in height as meteorological conditions alter and this variation causes differences in cosmic ray intensity at sea-level. Firstly, an increase in the mass of air between the layer and sea-level will result in greater absorption of the downcoming particles. Secondly, if the height at which mesons are produced is increased, the longer path will give the mesons greater opportunity to decay and fewer will reach sea-level. This increase in height can be caused by an increase in atmospheric pressure or by an increase in the "mean temperature" of that part of the atmosphere below the meson producing layer.

As meteorological conditions themselves alter with latitude, these two effects will produce a variation with latitude which is not due to any geomagnetic cause. Kupferberg(4) has discussed the "atmospheric latitude effect" produced by lati-



tudinal changes in height of the 100-mb. layer, and Rathgeber(5) has estimated that about one-third of the total latitude effect observed is due to a decrease in height of this layer experienced as the observer moves from the equator towards the poles.

It is customary to deduce pressure and temperature "coefficients" which, when applied, are intended to correct for the above two effects. Various methods of deducing such coefficients have been used. If there were a direct functional relation between sea-level temperatures and pressures and the height of the 100-mb. layer we would expect both coefficients to be negative and constant in their values. This, however, is not the case. Although a strong correlation no doubt exists, sea-level readings of temperature and pressure give unreliable information regarding the height of the meson-producing layer. This is probably the reason why coefficients for temperature and pressure deduced from sea-level data by different workers under different geographic and climatic conditions vary so widely.

It is therefore generally not possible to separate the "atmospheric latitude effect" from the "geomagnetic latitude effect" by making measurements at sea-level alone. However, where upper-air data from radiosonde balloons are obtained simultaneously with the cosmic ray results, reliable corrections can be made. This method has already been widely used for the analysis of observations recorded at a single locality(6-8). but when cosmic ray recordings are made on board ship it is generally impracticable to carry out radiosonde observations and one is forced to rely on measurements of temperature and pressure at sea-level. There is also the disadvantage that the ship may cross parallels of latitude quite rapidly, giving insufficient readings at any one latitude to reduce to a reasonable value the statistical errors for the readings representing that latitude.

Hence, the method to be used in deducing the correction factors for temperature and pressure constituted a problem. A further difficulty arose in the following fashion. Often, by correlating cosmic ray readings at sea-level with corresponding temperatures and pressures it is possible to deduce correction coefficients which satisfactorily smooth out large fluctuations caused by temperature and pressure changes. However, where latitude is changing it must be included in the analysis. Then, if the determination of regression coefficients by the method of *multiple* correlation between pressure, temperature, latitude, and cosmic ray intensity is attempted, one is faced with the complication that the variation of cosmic ray intensity with latitude is not linear.

The method finally adopted was as follows. Results at constant latitude were obtained at three sections of the voyage, namely, at Kure, the stay at Auckland, and the passage across the Tasman Sea. We could not combine all these latitude-constant readings, because the Kure readings treated separately gave coefficients widely different from those obtained from the Auckland and Tasman Sea sections. On the other hand, it was reasonable to combine the Auckland and Tasman Sea sections, as it was found that the Auckland and Tasman Sea coefficients, separately deduced, did not differ greatly from the combined coefficients. This might be expected from a consideration of the situation of Auckland in relation to the large surrounding water masses.



Another method was also tried. From cosmic ray readings which were obtained from count 3 at sea in the equatorial region we selected those whose corresponding pressures and temperatures lay within the ranges  $1013 \pm 5$  mb. and  $29 \pm 5^\circ\text{C}$ .

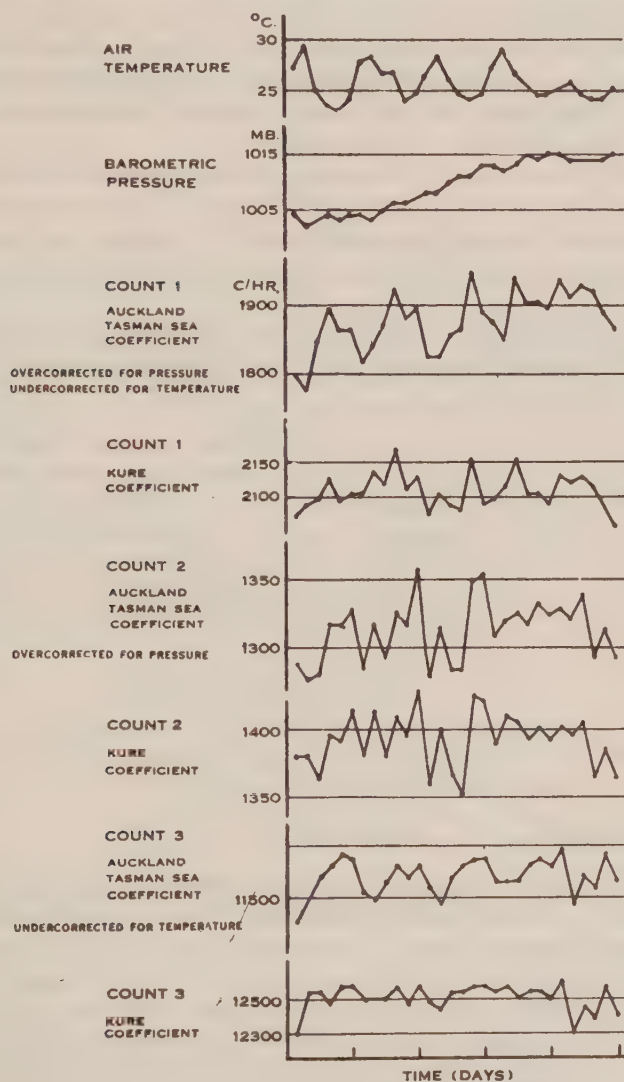


Fig. 3.—Four-hourly means of cosmic ray rates at Kure after correcting for pressure and temperature.

These readings were graphed against latitude. From a section of the graph where the gradient was nearly linear an approximate latitude correction factor was deduced and applied to the cosmic ray results over this section. The latitude-corrected results were then analysed to determine the multiple regression coefficients for

pressure and temperature. These agreed with the Auckland-Tasman Sea coefficients within the limits of the errors.

The temperatures and pressures used for the Kure period were obtained from the meteorological station at Hiroshima, 12 miles from Kure. This was necessary because regular observations aboard the *Duntroon* were not made at night while the ship was in port, and the thermograph and barograph were unreliable. However, there were no significant differences between the daytime readings on the *Duntroon*, the Hiroshima data, and records obtained from the nearby station of Iwakuni.

In Figure 3 are shown the four-hourly cosmic ray means obtained at Kure corrected for temperature and pressure (a) by using the Kure coefficients, and (b) by using the Auckland-Tasman Sea coefficients. The graphs for temperature and pressure at Kure show clearly that the second coefficients give inadequate correction.

Table 1 gives regression coefficients for temperature and barometric pressure deduced from the Japanese voyage and also, for comparison, those of the Antarctic(1).

TABLE 1  
PARTIAL REGRESSION COEFFICIENTS FOR COSMIC RAY-PRESSURE AND COSMIC RAY-TEMPERATURE

	Count 1	Count 2	Count 3
Kure ... ..	$b_{CP}^* + 0.015 \pm 0.058$ $b_{CT}^* - 0.540 \pm 0.139$	$+ 0.029 \pm 0.067$ $- 0.450 \pm 0.160$	$- 0.144 \pm 0.025$ $- 0.140 \pm 0.061$
Auckland—Tasman Sea	$b_{CP} - 0.204 \pm 0.042$ $b_{CT} + 0.036 \pm 0.103$	$- 0.064 \pm 0.037$ $- 0.148 \pm 0.092$	$- 0.161 \pm 0.030$ $+ 0.151 \pm 0.079$
Antarctic(1) Group 3 ...	$b_{CP} - 0.144 \pm 0.018$ $b_{CT} + 0.063 \pm 0.019$	$- 0.114 \pm 0.018$ $- 0.051 \pm 0.019$	$- 0.192 \pm 0.009$ $- 0.012 \pm 0.009$

\* $b_{CP}$ , % of mean/mb. ;  $b_{CT}$ , % of mean °C.

It is evident that the variation of the coefficients of this voyage from one count to the other and from one locality to the other does not agree with the variation which would be expected from the properties of the cosmic ray components. A closer scrutiny of the Auckland-Tasman Sea coefficients for counts 1 and 2 reveals that the variations tend to compensate each other. For instance, in count 1, a barometric coefficient which is too high balances a positive temperature coefficient which should be negative. It is suggested that these variations are due to circumstances similar to those prevailing in groups 1 and 2 of the Antarctic voyage(1). A correlation coefficient between temperature and pressure of  $-0.67$  demonstrates that these are not independent variables as is assumed in multiple correlation calculations. Under such conditions a physically significant variation is split up between the two variables in a ratio which depends largely on accidental features of their random fluctuations. Whether any physically significant differences in the temperature and barometric coefficients due to different meteorological conditions also exist is hidden by these fortuitous variations which greatly exceed them.

This argument, however, does not apply to the results from Kure, where the pressure *versus* temperature coefficient is only 0.23. Whilst it is realized that upper-air conditions of which we have no direct information may have caused anomalies in the regression coefficients for temperature and pressure, the following point is of interest. The temperature coefficient derived from the diurnal periods of cosmic ray intensity and of temperature (Table 3) agrees closely with the one obtained by multiple correlation (Table 1). It is feasible that there exists a diurnal cosmic ray variation not due to temperature but which is in phase with the diurnal temperature variation. If this be so it would contribute to the regression coefficient in the multiple correlation calculation.

(a) *Latitude Effect of Single Particles*

The four-hourly averages of the total narrow angle radiation (count 1), the hard narrow angle component (count 2), the soft narrow angle component (count 1 minus count 2), and the total wide angle radiation (count 3) are shown in Figure 4. The barometric pressure and temperature are also traced on the same graph.

From our knowledge of cosmic rays one would expect the barometric and temperature coefficients for the total radiation and the hard component to be not very different. For instance, the barometric coefficients for group 3, Antarctic voyage, are in the ratio  $0.144 : 0.114 = 1.26$ . In the Auckland-Tasman Sea section, however, the same coefficients stand in the ratio  $0.204 : 0.064 = 3.19$ . This, together with the difficulties outlined in the preceding Section, has led us to plot in Figure 4 the results as obtained without barometric and temperature corrections. The effect of applying these corrections is shown in Table 2.

TABLE 2  
MEASURED LATITUDE EFFECTS AND THEIR CORRECTIONS

	Count 1	Count 2	Count 1-2	Count 3	Count 4	Count 5
Latitude effect from Figures 4 and 5	$20 \pm 1$	$20 \pm 1$	$20 \pm 2$	$13 \pm 1$	$11 \pm 4$	$14 \pm 7$
Correction for resolving time ...					$-2 \pm 2$	$-5 \pm 4$
Barometric correction for 2.75 mb.	$-0.5$	$-0.2$		$-0.4$	$+2 \pm 1$	
Temperature correction for $-14.7^\circ\text{C}$ .	$+0.5$	$-2 \pm 1$		$+2 \pm 1$		
Corrected latitude effect(1%)... ..	$20 \pm 1$	$18 \pm 2$	$20 \pm 2$	$15 \pm 2$	$11 \pm 6$	$7 \pm 9$

The apparatus worked with only minor interruptions till  $9^\circ$  N. geomagnetic-latitude on the return journey. After replacement of one Geiger counter and adjustment of the counter voltage, count 2 gave results coinciding with those of the outward journey, but the readings of count 1 remained considerably higher till about  $40^\circ$  S. This part of the results was not used in the evaluation of the latitude effect.

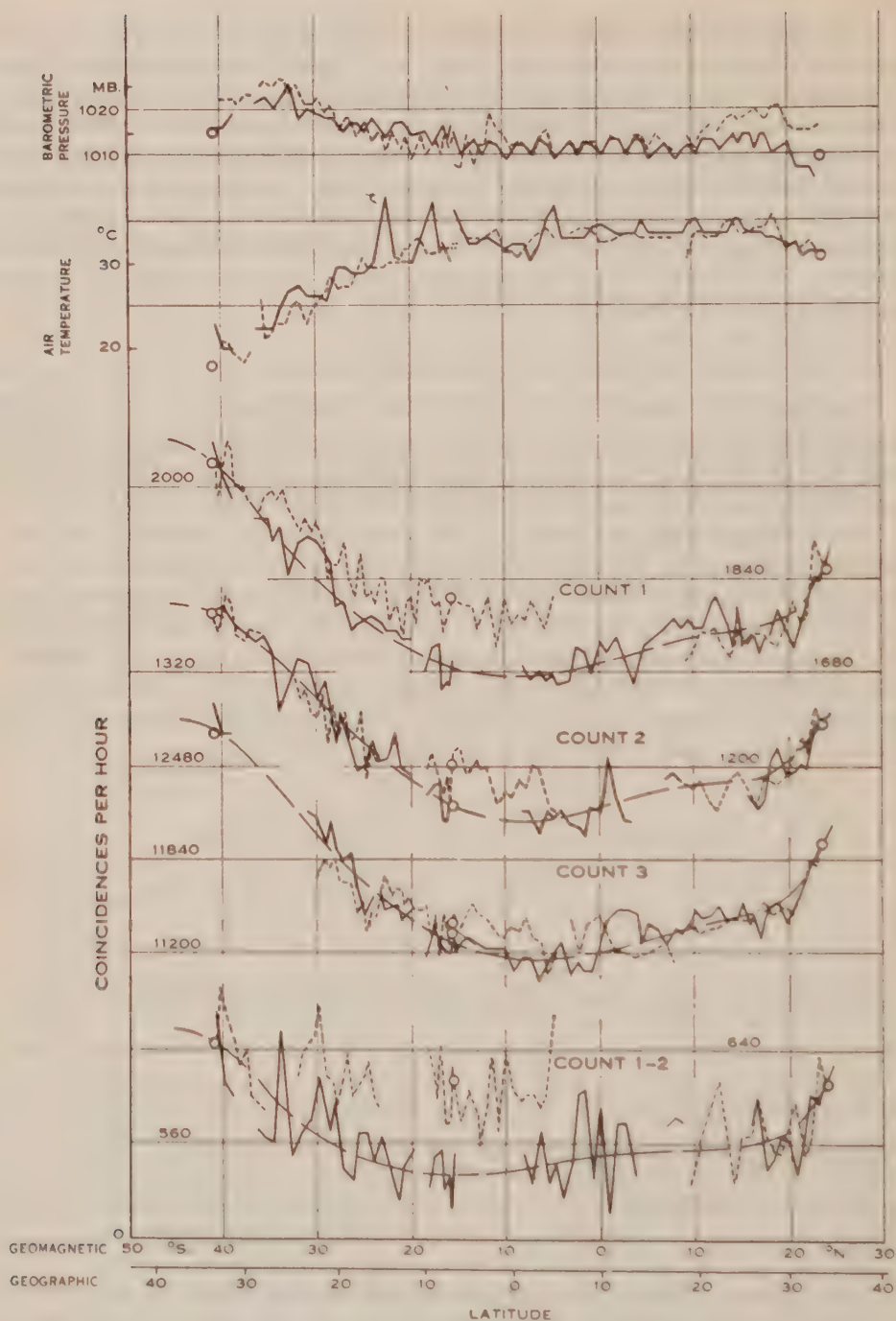


Fig. 4.—Uncorrected four-hourly means of pressure, temperature, and cosmic ray rates.

Count 1. Narrow aperture coincidences without filter.

Count 2. Narrow aperture coincidences filtered by 10 cm. lead.

Count 3. Wide aperture coincidences without filter.

— Outward to Japan.

.... Return from Japan.



Shortly after the resumption of counting the rate of counts 2 and 3 increased on July 28 ( $7^{\circ}$  S. geom.) by 3 and 2 per cent. respectively compared with the outward journey. They remained so till July 31. A backward extrapolation of the increased count rate of count 1 gives also a 3 per cent. excess for the same period. This increase in all three count rates is, as far as can be judged, not due to irregularities of the equipment, but may be due to an increase of the primary radiation.

It is also interesting to note that the curves for all three counts are not symmetrical and that a kind of plateau appears from  $10$  to  $20^{\circ}$  N. geom. As the recordings of Compton and Turner(9) in the Pacific Ocean show the same effect on some of their voyages only, it is assumed that the plateau is caused by meteorological factors.

During the stay at Kure the barometric pressure rose from 1004 to 1015 mb., to return gradually during the return trip to the value it had on the outward journey. One might have expected this to produce variations of 2.2, 0.7, and 1.8 per cent. for counts 1, 2, and 3 respectively. No variation of this magnitude is detectable in any of the curves. This, together with the abnormally small values of the barometric regression coefficients in Table 1, suggests that there was practically no barometric effect at Kure during our stay. This could be explained if the increase in pressure at sea-level were accompanied by a cooling of the higher atmosphere. If this were such as to cause a slight lowering of the meson-producing layer, then the two processes of absorption(10) and decay would result in no change in intensity at sea-level for changing barometric pressure.

In order to assess as accurately as possible the latitude effect, smooth curves were drawn through the graphs of the different counts. The extensions of these curves south of  $41^{\circ}$  S. geom. (Sydney) were deduced from ionization chamber measurements in the same region(11). They were obtained by applying corrections in the ratios of the latitude effect for the different counts to that of the ionization chamber measurements. As the extrapolation amounts to 2 per cent. at the most, the error introduced by this method is negligible. The latitude effect calculated from these compound curves is tabulated together with corrections for finite resolving time, barometric pressure, and sea-level air temperature (Table 2).

Comparison of counts 1 and 2 as well as of their difference, which is the intensity of rays unable to penetrate 10 cm. of lead, shows that the latitude effect for the total radiation, for the hard component and for the soft component, is the same within the errors of the experiment of 2 per cent. A calculation of the ratio of each four-hourly average of counts 1 and 2, which is not shown, confirms this result. Considering that practically all electrons are absorbed by 10 cm. of lead, this means that the latitude effect at sea-level is the same for the electron component as for the mesons.

It can be concluded that the latitude effect for the total vertical radiation uncorrected for sea-level temperature, pressure, or height of the meson-producing layer is  $20 \pm 1$  per cent. at sea-level. The latitude effect for radiation absorbed by 10 cm. of lead is the same but with an error of about 3 per cent. For the wide angle measurement the latitude effect amounts to only  $13 \pm 1$  per cent.

(b) *Latitude Effect of Extensive Showers*

The results of recordings of extensive showers during our voyage are shown in Figure 5. Count 4 gives the daily rate of showers in which the minimum separation of the two recorded rays is 0.8 m. and the maximum 1.2 m., count 5 has the same geometrical characteristics but in addition one of the rays has to penetrate through at least 10 cm. of lead. Averaging over intervals of 20 degrees of geomagnetic latitude yields latitude effects of  $11 \pm 4$  per cent. for the extensive showers and  $14 \pm 7$  per cent. for the penetrating showers.

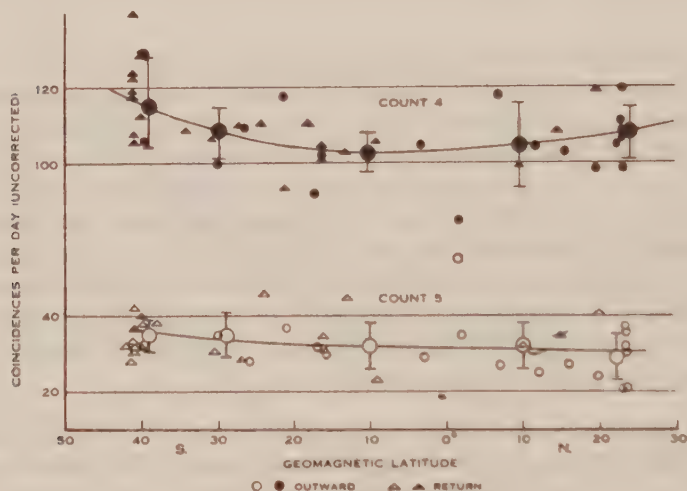


Fig. 5.—Daily rates of extensive showers.

Count 4. Showers of 1 m. spread without filter.

Count 5. Showers of 1 m. spread filtered by 10 cm. lead.

The rate of extensive showers is measured by the coincidences of counts 1 and 3 (Fig. 2), which are subject to a latitude effect. The accidental coincidences which are due to a finite resolving time,  $\tau$ , of the coincidence circuits and which are proportional to the product of the rates of counts 1 and 3, will show a latitude effect, too. Suppose that the accidental shower rate beyond the knee is  $2 N_1 N_3 \tau$ , then the rate at the equator becomes  $2 (1 - L_1) N_1 (1 - L_3) N_3 \tau$  in which  $L_1$  and  $L_3$  are the latitude effects of the pulse rates  $N_1$  and  $N_3$  of counts 1 and 3 beyond the knee. In the present case the latitude effect of the accidental coincidences is so large that it cannot be neglected. The corrections calculated for a resolving time of  $\tau = 26 \pm 3$  microseconds are given in Table 2. The error of the corrections, which is greater than the error of the resolving time, makes allowance for the fact that the resolving time was determined on a different set of equipment, but which was as similar as possible to the one of the voyage. A barometric coefficient of 0.7 per cent./mb.,

which was deduced from Auger and Daudin's measurements(12) was used to correct for barometric pressure.

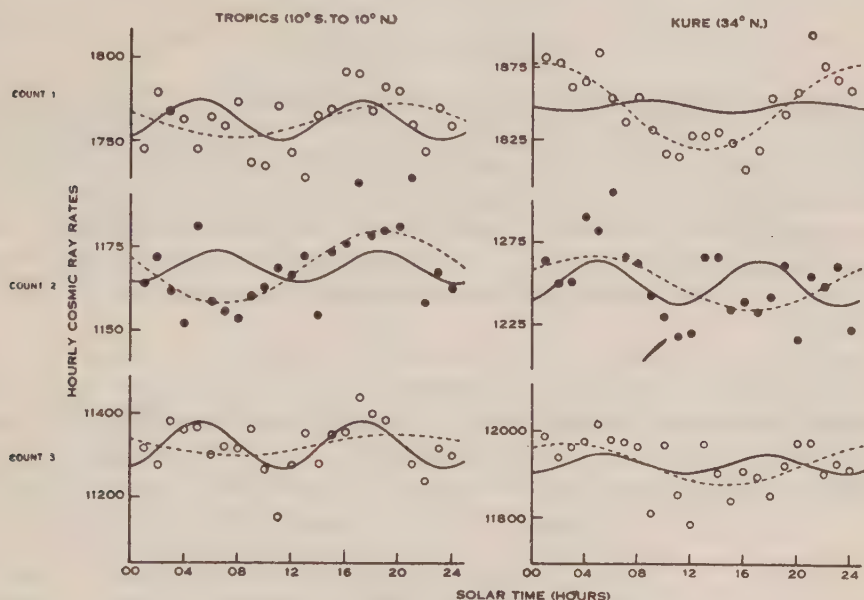


Fig. 6.—Diurnal periodicities.

----- Diurnal period.  
 — Semi-diurnal period.

To make the best use of all available information we also subjected the published results of Neher and Pickering(13) as well as count 4 to a multiple correlation calculation for latitude and pressure. Whilst we realize that the latitude effect is not a linear function of geomagnetic latitude and should therefore not be calculated by linear multiple correlation, we believe that the errors introduced in this way are smaller than the statistical errors. The latitude effects obtained by multiplying the regression coefficients by the difference in latitude over which the measurements extended amounted to  $8 \pm 4$  for the data of Neher and Pickering(13) and to  $13 \pm 6$  per cent. for count 4. This latter reduces to  $11 \pm 6$  per cent. after correction for accidental coincidences. The barometric regression coefficients were found to be  $1.9 \pm 1.0$  and  $0.1 \pm 0.3$  per cent./mb. respectively. The first agrees with other measurements(12). We attribute the smallness of the second to the nearly constant pressure during our voyage. The small variations in the shower rate due to pressure variations were much smaller than the random fluctuations and thus disappeared in the limited statistical material.

The combination of the two results makes it certain that a latitude effect of some 10 per cent. exists for extensive showers of 1 m. base.



Considering the large error of the latitude effect found for penetrating extensive showers (count 5),  $7 \pm 9$  per cent., the only safe inference is that the existence of a latitude effect is not excluded, but that its magnitude should not exceed 20 per cent.

(c) *Diurnal Variations*

The curves of Figure 4 show that in equatorial regions, there is a tendency for a maximum in the cosmic ray intensity to occur in the afternoon. To check this, uncorrected hourly readings were chosen from both the forward and return journeys over the section  $10^\circ$  S. to  $10^\circ$  N. geographic latitude in such a manner that there was no apparent bias towards any particular hour due to the latitude effect or the higher count rate on the return voyage. These were subjected to a harmonic analysis. The results of this and of an analysis of the recordings at Kure are shown in Figure 6 and Table 3.

TABLE 3  
CALCULATED PERIODICITIES AND TEMPERATURE AND BAROMETRIC COEFFICIENTS DERIVED FROM THE PERIODICITIES

	Count 1		Count 2		Count 3	
	Amplitude (%)	Time of Maximum Hours	Amplitude (%)	Time of Maximum Hours	Amplitude (%)	Time of Maximum Hours
<i>Tropics</i> ( $10^\circ$ S.— $10^\circ$ N.)						
Diurnal ... ..	$0.6 \pm 0.4$	$19.6 \pm 2.6$	$0.9 \pm 0.4$	$18.9 \pm 1.7$	$0.2 \pm 0.1$	$20.0 \pm 2.1$
Temperature coeff. (%/°C.) ... ..	$-0.5 \pm 0.3$		$-0.7 \pm 0.3$		$-0.2 \pm 0.1$	
Semi-diurnal ... ..	$0.7 \pm 0.4$	$17.2 \pm 2.4$	$0.4 \pm 0.4$	$18.5 \pm 3.8$	$0.5 \pm 0.1$	$17.3 \pm 1.0$
Barometric coeff. (%/mb.) ... ..	$-0.5 \pm 0.3$		$-0.3 \pm 0.3$		$-0.4 \pm 0.1$	
<i>Kure</i> ( $34^\circ$ N.)						
Diurnal ... ..	$1.4 \pm 0.3$	$0.9 \pm 0.9$	$1.3 \pm 0.4$	$4.4 \pm 1.1$	$0.4 \pm 0.2$	$2.6 \pm 1.5$
Temperature coeff. (%/°C.) ... ..	$-0.6 \pm 0.1$		$-0.5 \pm 0.2$		$-0.2 \pm 0.1$	
Semi-diurnal ... ..	$0.2 \pm 0.3$	$20.8 \pm 7.8$	$1.1 \pm 0.4$	$17.0 \pm 1.2$	$0.2 \pm 0.2$	$17.7 \pm 3.4$
Barometric coeff. (%/mb.) ... ..	$-0.2 \pm 0.5$		$-1.5 \pm 0.5$		$-0.3 \pm 0.3$	

Despite the short series of results available (nine days in the tropics and five days at Kure) the semi-diurnal period(14) appears quite clearly and is nearly exactly opposed in phase to the barometric semi-diurnal oscillation of the atmosphere(15). Although the amplitude cannot be considered significant except for count 3 in the tropics, the identity of the phase in all measurements within their standard errors shows that the effect exists.



Taking as the amplitude of the semi-diurnal pressure wave 1.25 mb. in the tropics and 0.70 at Kure(15), corresponding barometric coefficients were obtained. They are, on the average, about twice as great as the coefficients for non-periodic pressure variations. The present evidence confirms that of other measurements(14). It must be concluded that the mechanism linking the height or heights of meson production with the variations of pressure at sea-level differs in the two cases.

The maxima of the diurnal period coincide in the tropics and at Kure with the daily temperature minima. Considering the general variability of the temperature coefficient with geographical conditions the agreement between the results at Kure and in the tropics, which were derived from the amplitudes of the diurnal period of cosmic ray intensity and of the temperature cycle, is rather remarkable.

#### IV. ACKNOWLEDGMENTS

The cost of the experiment was covered by a grant from the A.N.A.R. Expedition. The authors wish to express their thanks to Professor L. H. Martin for his constant interest and to D. E. Caro, J. S. Grice, and C. L. Hamblin for their work in redesigning the electronic equipment, and to the United States Air Force Meteorological Services and the Central Meteorological Observatory, Tokyo, for supplying meteorological data from a number of stations.

#### V. REFERENCES

- (1) CARO, D. E., LAW, P. G., and RATHGEBER, H. D.—*Aust. J. Sci. Res.* A 1: 261 (1948).
- (2) JÁNOSSY, L.—“Cosmic Rays.” (Oxford, 1948.)
- (3) MONTGOMERY, D. J. X.—“Cosmic Ray Physics.” (Princeton, 1949.)
- (4) KUPFERBERG, K. M.—*Phys. Rev.* 73: 804 (1948).
- (5) RATHGEBER, H. D.—*Nature* 162: 303 (1948).
- (6) DUPIERIER, A.—*Proc. Phys. Soc.* 57: 464 (1945).
- (7) DUPIERIER, A.—*Ibid.* 61: 34 (1948).
- (8) BEARDSLEY, N. F.—*Phys. Rev.* 59: 233 (1941).
- (9) COMPTON, A. H., and TURNER, R. N.—*Ibid.* 52: 799 (1937).
- (10) DUPIERIER, A.—*Terr. Magn. Atmos. Elect.* 49: 1 (1944).
- (11) GILL, P. S.—*Phys. Rev.* 55: 1151 (1939).
- (12) AUGER, P., and DAUDIN, J.—*Ibid.* 61: 91 (1942).
- (13) NEHER, H. V., and PICKERING, W. H.—*Ibid.* 58: 665 (1940).
- (14) NICHOLSON, P., and SARABHAI, V.—*Proc. Phys. Soc.* 60: 509 (1948).
- (15) HANN, J., von (Editor, R. Suring).—“Lehrbuch der Meteorologie.” Vol. 5. (Tauchnitz: Leipzig, 1939.)

# MEASUREMENTS OF SOLAR RADIATION AT A WAVELENGTH OF 50 CENTIMETRES DURING THE ECLIPSE OF NOVEMBER 1, 1948

By W. N. CHRISTIANSEN\*, D. E. YABSLEY\*, and B. Y. MILLS\*

(Plates 1-2)

[*Manuscript received July 27, 1949*]

## Summary

Radio-frequency power received from the sun at a wavelength of 50 cm. was measured at three well-separated places during the solar eclipse of November 1, 1948. Abrupt changes in slope on the records of received flux density were interpreted as being the result of the covering and uncovering on the sun of small areas of great radio brightness. These areas were found to be associated with some visible sunspots, with positions previously occupied by sunspots, and with one prominence. The average effective temperature of the bright areas was about  $5 \times 10^6$  °K., and the areas contributed a total power of roughly one-fifth of that from the entire sun.

After the effects of active areas had been taken into account, the remaining four-fifths of the power received from the sun was found to originate from a source larger than the visible disk. About 40 per cent. of the power from this source originated outside the edge of the visible disk. The results were consistent with a theoretical distribution of brightness on the source, which involved limb-brightening.

The relative magnitudes of the two circularly-polarized components of the solar radiation showed small differences as the bright areas were eclipsed. No predominance of one component was seen when one hemisphere of the sun was eclipsed; hence no effects of any general magnetic field on the sun were detected.

## I. INTRODUCTION

In the last few years the characteristics of solar radiation at centimetre and metre wavelengths have been studied intensively, but few observations have been made at decimetre wavelengths. Reber(1), observing at 62.5 cm. found that the received intensity corresponded to an apparent solar temperature of  $10^6$  °K., with small day-to-day variations closely corresponding with variations in the area of sunspots. Lehany and Yabsley(2) reported similar characteristics on 25 and 50 cm., where the apparent temperatures varied between  $0.1$  and  $0.2 \times 10^6$  °K., and  $0.5$  and  $1.0 \times 10^6$  °K., on the respective wavelengths, and correlated closely with sunspot area.

In addition to the small day-to-day variations, sudden large-scale increases in intensity were observed on rare occasions. The mechanism of origin of these is unknown, but it seems unlikely that it is thermal. Pawsey and Yabsley(3) showed that the relatively steady radiation from the sun, at wavelengths of 50 and 25 cm., involved two components. One component was postulated as being the thermal radiation from the "quiet" sun, while the other was intimately connected with sunspots. An analysis of measurements of radiation intensity at the two wavelengths

\*Division of Radiophysics, C.S.I.R.O.

showed that the apparent temperature of the sun in the absence of sunspots was  $0.5 \times 10^6$  °K. at 50 cm. and  $0.1 \times 10^6$  °K. at 25 cm. wavelengths. These temperatures lie between the apparent temperatures of about  $10^6$  °K. observed at metre wavelengths and about  $10^4$  °K. observed at centimetre wavelengths. The change in apparent temperature with wavelength has been explained by Ginsburg(4) and Martyn(5) on the assumption that on the sun the corona is at a very much higher temperature than are the chromosphere and photosphere. At metre wavelengths the radiation originates chiefly in the corona, while at centimetre wavelengths it originates mainly in the chromosphere. Decimetre-wavelength radiation occupies an intermediate position. Martyn concluded that at decimetre and centimetre wavelengths radiation from the vicinity of the limb of the sun would originate in higher temperature regions of the solar atmosphere than would radiation from the centre of the disk. Hence at these wavelengths, and particularly at decimetre wavelengths, marked "limb-brightening" was predicted.

Waldmeier and Müller(6) and Smerd(7) have made a more detailed estimate of the magnitude of this effect, including a prediction of the apparent radius of the sun at different wavelengths.

The only published observations relating to this effect are those of Covington(8) on 10.7 cm. and of Sander(9) on 3.2 cm. wavelengths. Both obtained records of the total intensity of solar radiation during partial eclipses of the sun. Sander's observations showed some evidence for the existence of limb-brightening, but the most definite conclusion to be drawn from these observations was that at 10.7 cm. the radiation from the vicinity of a large sunspot group located near the centre of the disk was considerably more intense than that from the remainder of the disk.

The solar eclipse of November 1, 1948, which was visible as a partial eclipse over most of the Australian continent, provided an opportunity both to investigate the distribution of the radiation "brightness" over the solar disk and also to determine the apparent size of the sun. The wavelength chosen for the observations was 50 cm.\* and observations were taken at three well-separated locations where the circumstances of the eclipse were markedly different. This procedure was adopted in order to make it easier to distinguish between the three possible anticipated effects, namely, increased size of disk, limb-brightening, and increased brightness near sunspots.

Considerable interest also attaches to the question of polarization of solar radio-frequency radiation. It has been known for some time that the "enhanced radiation" from the region around sunspots exhibits circular polarization at metre wavelengths, and there is some evidence that similar effects occur at shorter wavelengths. It has also been predicted by Martyn that the existence of a general magnetic field on the sun should cause the thermal radiation from each hemisphere to be partially circularly-polarized (in opposite senses in the two hemispheres, the total radiation being then random). As the circumstances of the present eclipse were such that at maximum phase the portion of the sun exposed was either wholly or predominantly the northern hemisphere at the observing stations, there was an opportunity to

\*Observations were made also at a wavelength of 10 cm. by J. H. Piddington and J. V. Hindman, and at 3.18 cm. by H. C. Minnett and N. R. Labrum. Their observations are being published separately.



observe this effect as well. Accordingly, at two of the stations it was arranged to compare the intensities of the left- and right-handed circularly-polarized components of the sun's radiation during the course of the eclipse.

## II. EXPERIMENTAL PROCEDURE

The three stations chosen for observing the eclipse were Sydney in New South Wales, Rockbank near Melbourne, Victoria, and Strahan in Tasmania. The circumstances of the eclipse at each place are summarized in Table 1.

TABLE 1

Station	Magnitude	Eastern Australian Standard Time			
		Beginning of Eclipse	Greatest Phase	End of Eclipse	Sunset
Sydney ...	0.63	1647	1745	—	1824
Rockbank	0.72	1639	1741	1838	1854
Strahan ...	0.85	1634	1738	1837	1901

The receiving equipment used at Sydney was the same as that described by Lehany and Yabsley(2) except that the 18 by 16 ft. paraboloid aerial was mounted on a polar axis and was arranged with a motor drive to follow the sun. This aerial was linearly polarized.

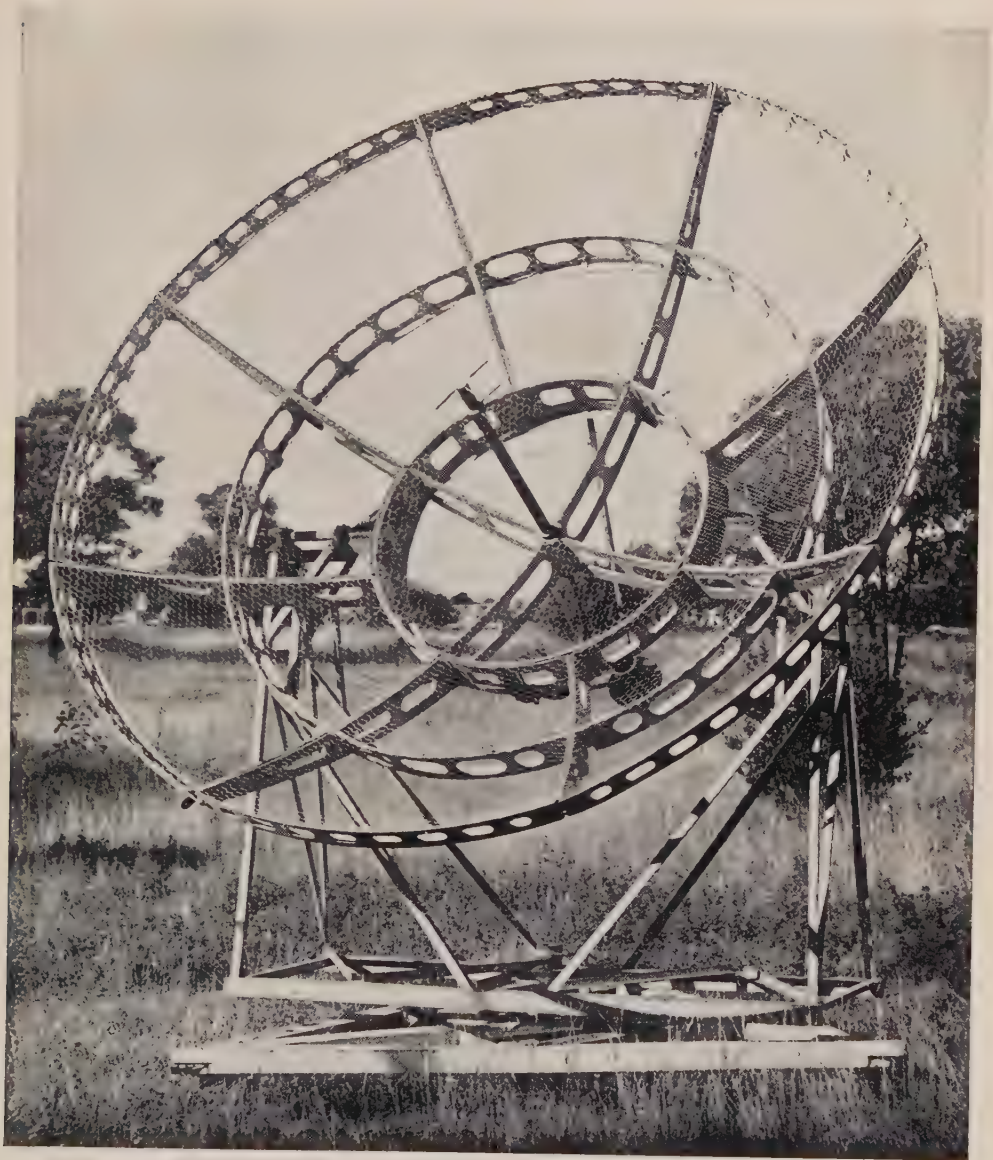
At both Rockbank and Strahan a paraboloid reflector having a diameter of 10 ft., and mounted so as to be rotatable in azimuth and elevation, was used (Plate 1). The feed consisted of crossed dipoles with reflectors. A coaxial transmission line was taken from each dipole, the two lines being connected in parallel before reaching the receiver. When one line was a quarter wavelength longer than the other, the aerial accepted circular polarization of one sense. The sense of polarization was reversed by changing the length of the other line by a half wavelength. Adjustment of the line lengths to make the aerial circularly polarized, and to reverse the sense of polarization, was achieved by means of telescopic pieces of coaxial line.

The receivers at all stations consisted of a quarter-wave transmission-type cavity-resonator, followed by a conventional crystal converter, 30 Mc. s. intermediate-frequency amplifier, and diode second detector. The resonator was necessary to exclude energy at the image frequency and also at higher frequency bands, which could be converted to the intermediate frequency by harmonics of the local oscillator. After rectification, the signals passed into a D.C. amplifier which was connected to a recording milliammeter.

With the small aeriels used at Rockbank and Strahan, the ratio of signal power from the uneclipsed sun to the internal noise power at the receiver input was between 0.10 and 0.15, so that great care was required for accurate measurement of the power flux density from the sun during the course of the eclipse.

Stability of receiver gain depends, apart from good components and absence of feedback, on constancy of temperature and stability of supply voltages. These were





CHRISTIANSEN, YABSLEY, AND MILLS.— MEASUREMENTS OF SOLAR RADIATION AT  
A WAVELENGTH OF 50 CENTIMETRES DURING THE ECLIPSE OF NOVEMBER 1, 1948



achieved by running the receivers for some hours before the eclipse began, and by operating all power supplies from rapidly-acting electronic voltage stabilizers. Errors due to the residual small drift of receiver gain (about 1 per cent. per hour) were eliminated by turning the aerial away from the sun every few minutes during the course of the eclipse to check the receiver gain.

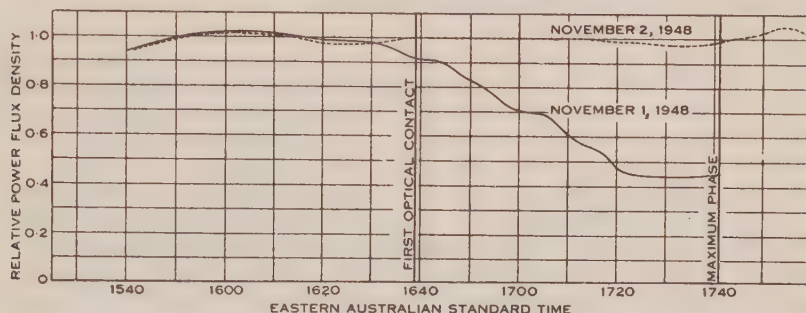


Fig. 1.—Records obtained at Rockbank on the day of the eclipse, and on the following day.

At Sydney, where the large aerial was used, the ratio of signal power from the sun to the internal noise power of the receiver was approximately four times the value at the other stations. Stability requirements were therefore considerably less stringent and it was found sufficient for the receiver gain to be checked by turning the aerial away from the sun only at the beginning and end of the record.

The aerials at Rockbank and Strahan were rotated by hand, and tables showing elevation and azimuth of the sun were prepared against the possibility of cloudy weather during the eclipse. At both stations, the accuracy of alignment of the aerial was checked on days previous to the eclipse by comparing the position of the aerial which gave maximum output with the calculated position of the sun. During the eclipse the aerial was adjusted every few minutes and the probable error in solar intensity measurements due to misalignment was estimated as 1 per cent.

Another possible source of error, and the principal one at Sydney, was that resulting from energy reflected from the ground being received by the aerials. This was possible because the eclipse took place late in the afternoon when the sun's elevation was small. The elevation of the sun was lowest at Sydney, but as a partial compensation for this the larger aerial used there was more highly directive than the smaller ones used at Rockbank and Strahan. (The beam width was  $7^\circ$  for the larger aerial against about  $15^\circ$  for the other two.) The effect of ground reflections was determined from records of power received from the sun during the late afternoon on one or two days following the eclipse. A typical record (made at Rockbank) is shown in Figure 1, and, for comparison, an eclipse record is also displayed. The scales used on the two records are arbitrary, so that the departure from the mean is the significant feature of the records. Each record shows a deviation of about 2.5 per cent. from the mean in the half-hour before the time of the eclipse, but follow each other very closely. The agreement between the records is such that, after the

correction has been added to the eclipse record, the error expected to result from ground reflections is about 1 per cent.

The records for Strahan are similar to the one shown in Figure 1. At Sydney the ground effect was more marked and larger correction factors were involved. A comparison of two days' records on the uneclipsed sun suggests that the error in the final result is probably less than 3 per cent.

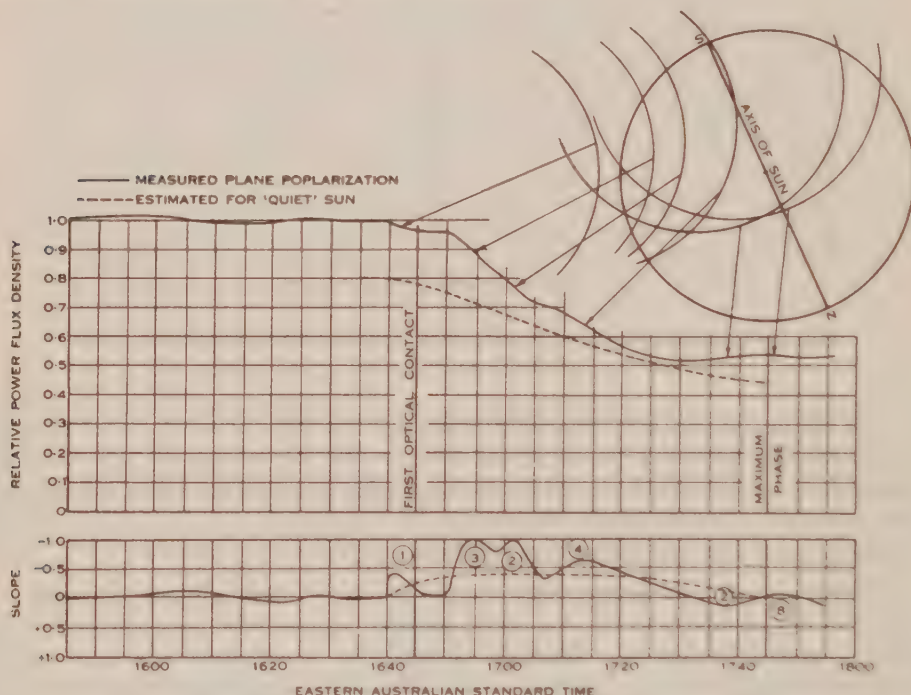


Fig. 2. —The eclipse record (corrected) obtained at Sydney. The circular arcs show the positions of the moon at the times of occurrence of the peaks in the slope curves.

### III. EXPERIMENTAL RESULTS

On the day of the eclipse, the received flux density at 50 cm. wavelength appeared to be steady. Several small groups of spots were visible on the solar disk (Plate 2). The total area of the spots, relative to the area of the visible disk of the sun was  $85 \times 10^{-5}$ , which is a small value at the present phase of the sunspot cycle.

At each of the three locations, recording of radiation intensity commenced several hours before the eclipse and continued until the end of the eclipse, except at Sydney, where sunset occurred before the last contact. At Sydney and Melbourne good records were obtained only until maximum phase of the eclipse. At Sydney large disturbances resulting from ground reflections made the record unreliable after this time, while at Melbourne a temporary disturbance in the power supply occurred shortly after the maximum phase of the eclipse. At Strahan, ground



reflection effects were small until 20 minutes after the maximum phase of the eclipse but increased rapidly thereafter, so the records have been used only to this time.

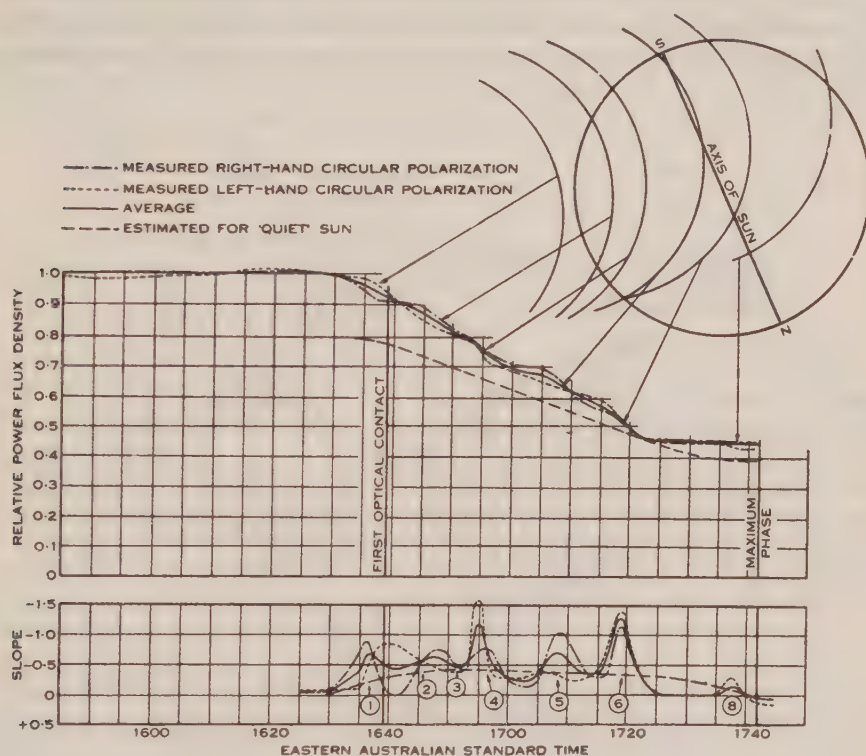


Fig. 3.—The eclipse record (corrected) obtained at Rockbank. The circular arcs show the positions of the moon at the times of occurrence of the peaks in the slope curves.

The records obtained at Strahan and Rockbank are discontinuous because the aerials were directed away from the sun at frequent intervals for the purpose of receiver calibration. In addition, the sense of polarization of the aerials was changed periodically, so that the record for each sense of polarization is discontinuous. The intermediate values were then interpolated. No serious error is likely to be caused by this procedure because the changes were made at close intervals.

The record at Strahan after 1649 hours is for right-hand circularly-polarized radiation only, because a mechanical fault in the equipment prevented changes in the sense of polarization after this time.

The relative magnitudes of the radio-frequency power flux density from the sun during the course of the eclipse at Sydney, Rockbank, and Strahan, respectively, are seen in Figures 2, 3, and 4. These curves have been corrected for ground-reflection effects and also for non-linearity in the relation between the power input

to the receiver and the increment of current flowing in the recording milliammeter.

In each figure, the positions of the optical disk of the moon relative to that of the sun are shown for various phases of the eclipse. These positions were calculated from astronomical data. The calculations of the circumstances of the eclipse at Sydney were checked by means of a series of accurately timed photographs; at

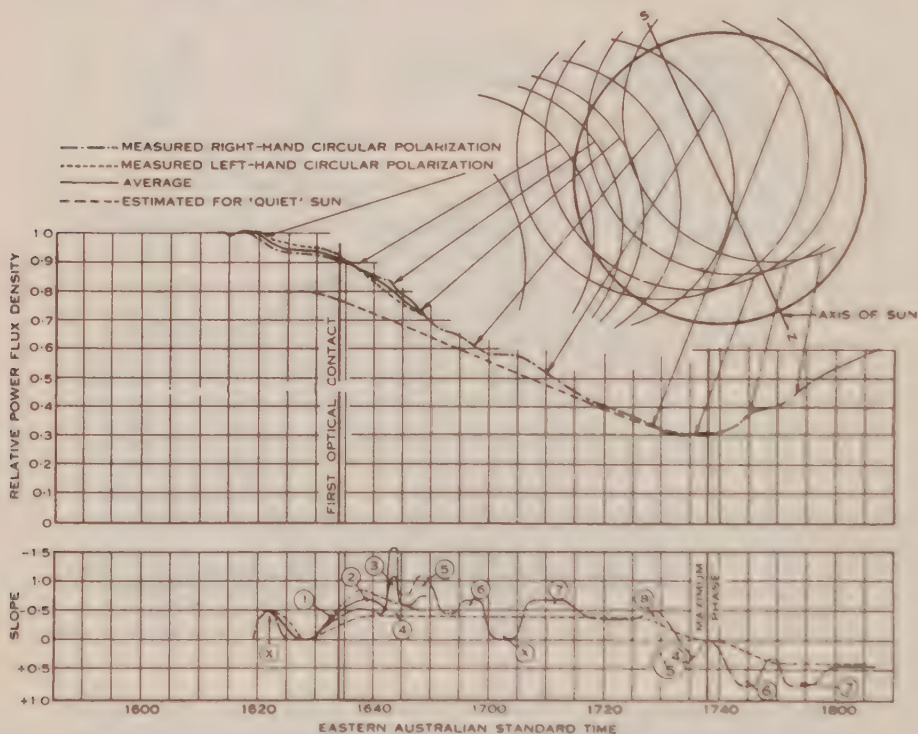


Fig. 4.—The eclipse record (corrected) obtained at Strahan. The circular arcs show the positions of the moon at the times of occurrence of the peaks in the slope curves.

Melbourne a single photographic check was made during a brief period when the sun was visible.

A curve marked "estimated for the 'quiet' sun" is also shown in each of the three figures. The basis of these estimations will be dealt with in Subsection IV (b).

#### IV. ANALYSIS OF RESULTS

##### (a) *The Existence and Location of Small "Bright" Areas on the Solar Disk*

Each of the curves of power flux density plotted against time during the eclipse shows, over short periods, changes of slope that are much greater than probable instrumental fluctuations. These changes are seen much more clearly in the lower

curves of Figures 2, 3, and 4 which give the rate of change of observed flux density with time. Possible explanations of these changes in slope are that (a) they result from changes in the radiation brightness of the sun, or (b) they are caused by the obscuring of small bright areas on the solar disk by the moon. The first possibility may be eliminated, because these changes did not occur simultaneously at the three observing stations. The second explanation appears to be the correct one and its correctness may be tested by examining the consistency of the three records.

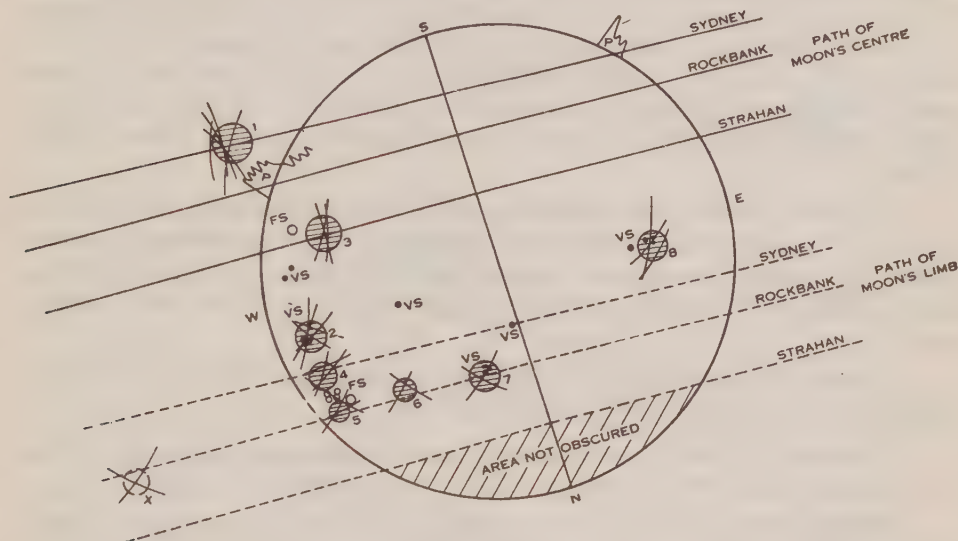


Fig. 5.—The location of small bright areas on the solar disk, and their relation to some optical features. Bright areas are indicated by numerals.

V.S., Visible sunspot groups; F.S., Position of sunspots 27 days before the eclipse; P., Prominences.

If the three records are consistent, it should be possible to locate the positions of the bright areas by triangulation. This was done as follows:

The times at which the rate of change of power flux density showed maximum values were obtained from the lower curves in Figures 2, 3, and 4. On a diagram of the solar disk, circles were drawn showing the edge of the obscured area at these times. It would be expected that the arcs would form triple intersections, if the postulate regarding the origin of the irregularities on the radiation-intensity/time curves is correct. Such triple intersections are shown in Figure 5. It should be noted that the number of irregularities in the power flux-density curves is greatest for Strahan and least for Sydney. This is not unreasonable, because the part of the solar disk obscured was greatest and least at these two places. Hence the number of such triple intersections will be limited to those in which the Sydney record is concerned.

Bright areas could also be identified when they were uncovered later in the eclipse because the rate-of-change curves in Figures 2, 3, and 4 had maximum positive values at these times. Circular arcs representing the position of the edge of the



moon at such times were drawn and are included in Figure 5. Thus some of the bright areas have more than three intersections associated with them.

With the large number of intersections involved, the elimination of ambiguities in the interpretation was not easy. The most consistent interpretation involved the assumption that the second irregularity on the Strahan record corresponded with the first irregularities on the other records. The first Strahan irregularity was then attributed either to a sudden fluctuation in receiver gain or to the obscuring of a source of radiation (with a large right-hand circularly-polarized component) well removed from the optical limb of the sun and outside the apparent path of the moon in the Rockbank and Sydney observations. The latter explanation is made plausible because such a source was apparently uncovered at Strahan later in the eclipse when the right-hand component remained stationary for a period of five minutes.

The times of covering and uncovering of the bright areas, which are numbered from 1 to 8 on Figure 5, are indicated by numbers on the rate of change curves in Figures 2, 3, and 4. The symbol *X* marks the area in Figure 5 and the two slope curve irregularities in Figure 4 corresponding to the first disturbance on the Strahan record.

An attempt to identify bright areas with other physical features of the solar atmosphere produced\* the following results:

(a) Three of the bright areas were situated around sunspot groups. (Three other sunspot groups were present, but were not associated with a bright area.)

(b) Three of the bright areas were close to the positions occupied by sunspot groups 27 days previously. The positions of these groups are indicated by small open circles in Figure 5. Allen(10) has shown that faculae, Ca flocculi, and H $\alpha$  flocculi tend to persist after the disappearance of sunspots with which they are related. The existence of bright areas in the vicinity of places previously occupied by sunspot groups suggests that a similar effect may be present here.

(c) At least one bright area is seen to exist outside the visible disk of the sun. In the case of No. 1 the distance from the centre of the sun is roughly 1.25 the optical radius of the solar disk, or at least  $1.7 \times 10^5$  km. above the photosphere. At this position a large stable prominence was seen on the day of the eclipse, by members of the staff of the Commonwealth Observatory at Mt. Stromlo. A second prominence also was observed on this day, but there is no evidence from the eclipse records that it had any pronounced radio-brightness.

(d) An intersection appears outside the north-west quadrant of the solar disk. The intersection, mentioned earlier, occurs at a place which was eclipsed only at Strahan, and in the absence of confirmatory evidence, must be considered doubtful.

\*Since the manuscript of this paper was prepared, Dr. S. B. Nicholson, of the Mount Wilson Observatory, has kindly sent to the authors detailed results of observations of the optical features of the sun on the day of the eclipse. No correspondence in detail, additional to that mentioned in the text, was found between optical features and areas of high radio-brightness on the solar disk. In general, however, the areas of high radio-brightness were found in the part of the solar disk where H $\alpha$  markings or prominences were most apparent.



The position indicated is at a distance of 1.8 times the optical radius of the solar disk from the centre of the sun.

(b) *The General Distribution of Radiation Brightness over the Disk of the Sun*

Since the areas of great brightness which have already been identified appeared to be quite limited in their extent, it is natural to consider the radiation from them

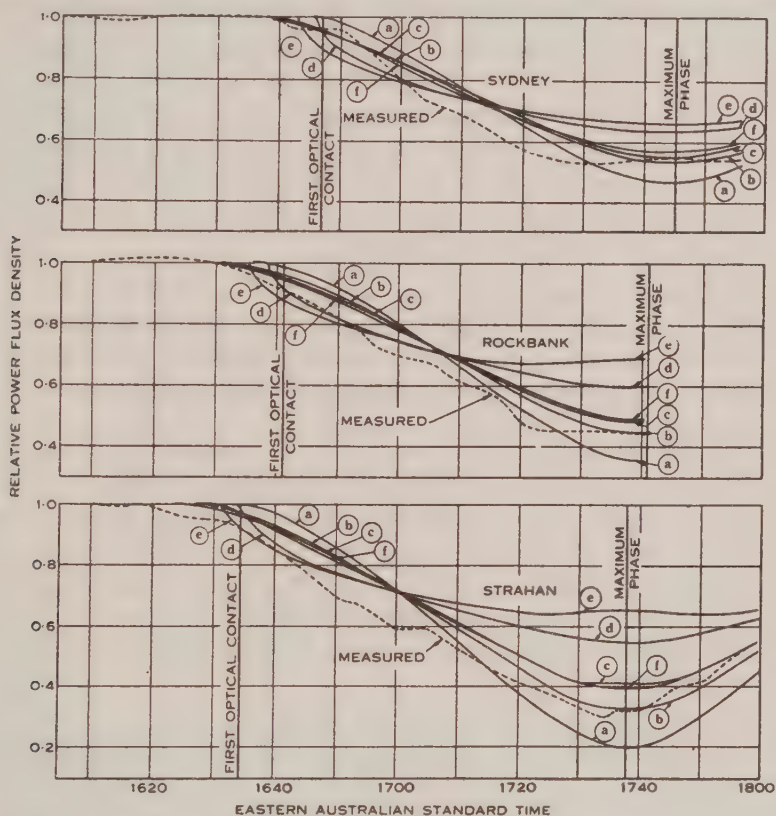


Fig. 6.—Curves showing the effect at each station of eclipsing artificial sources with the following distributions of brightness :

- (a) Uniform over the visible disk of the sun.
- (b) Uniform over a disk with a radius 1.2 times that of the visible disk.
- (c) Uniform over a disk with a radius 1.3 times that of the visible disk.
- (d) Uniform over a narrow circular ring with a radius equal to that of the visible disk.
- (e) Uniform over a narrow ring with a radius 1.1 times that of the visible disk.
- (f) Having 0.7 of its energy in a source of type (b) and the remainder in a source of type (d). (This approximates to a theoretically derived distribution.)

The dotted lines indicate the experimental curves.

to be additional to that from a source with circular symmetry with respect to the centre of the solar disk. This section is concerned with the isolation of such a source and the determination of the form of its distribution.

In Figure 6, in addition to curves representing recorded values of flux density plotted against time, various calculated relations are plotted. These show the manner in which the received flux density would have changed at each of the three observing points had the brightness of the disk been as follows:

- (a) Uniform over the visible disk of the sun.
- (b) Uniform over a disk with a radius 1.2 times that of the visible disk.
- (c) Uniform over a disk with a radius 1.3 times that of the visible disk.
- (d) Uniform over a narrow circular ring with a radius equal to that of the visible disk.
- (e) Uniform over a narrow ring with a radius 1.1 times that of the visible disk.
- (f) Having 0.7 of its energy in a source of type (b) and 0.3 of its energy in a source (d). This composite source is an approximation to the brightness distribution derived theoretically by Smerd(7) which is in rough agreement with that of Waldmeier and Müller(6).

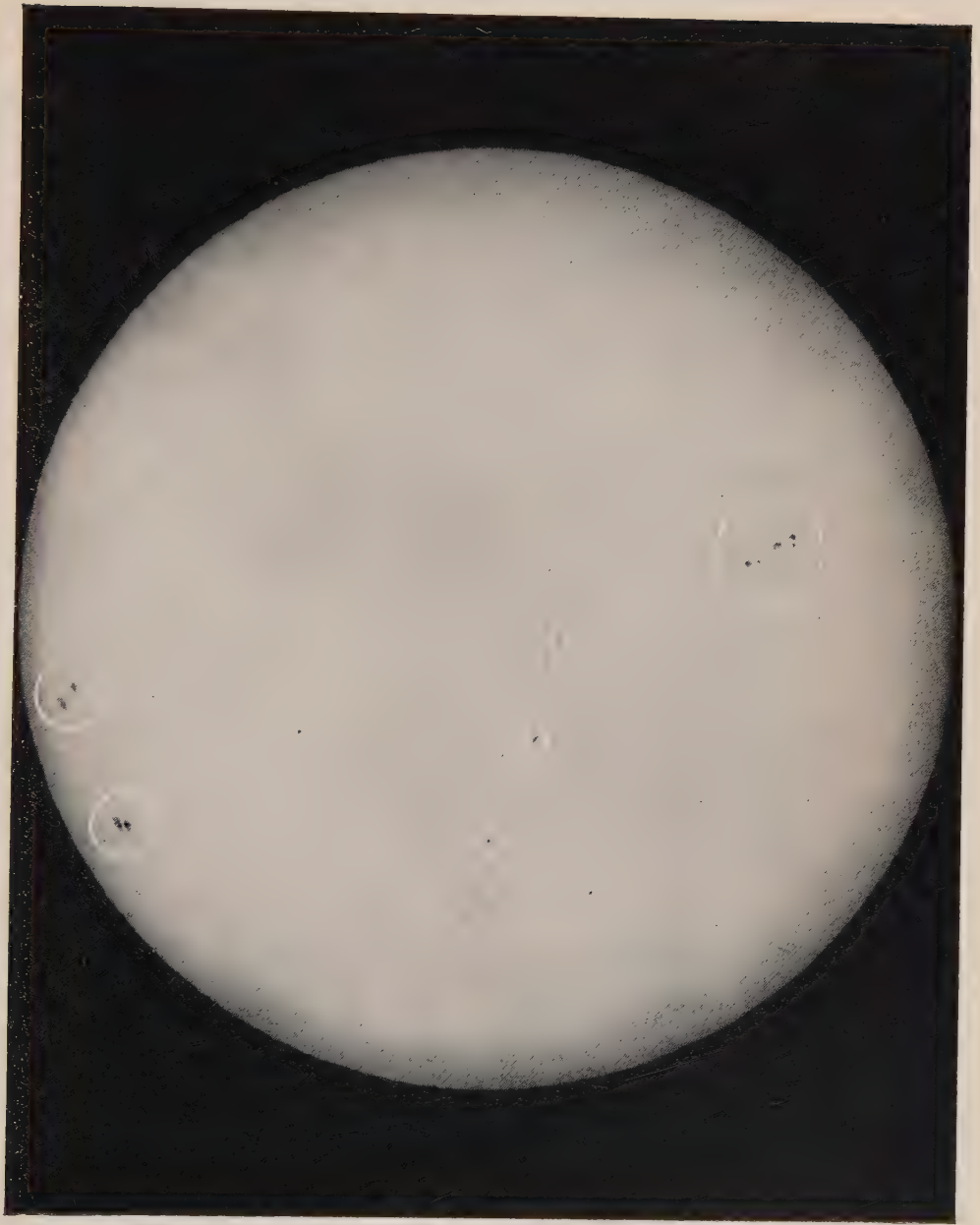
It will be seen in Figure 6 that at one particular phase of the eclipse, the proportional decrease in power flux density from the sun is practically the same for all the types of source considered. This occurs when the fraction of the area of the visible disk of the sun obscured by the moon is equal to the fraction of the periphery obscured.

It is noteworthy that at such times the decrease in power flux density as measured at each location is greater than the decrease indicated by any of the six theoretical distributions or, in fact, by any distribution circularly symmetrical with respect to the centre of the optical disk of the sun. This is to be expected since several of the small bright areas had already been eclipsed. The differences between the observed and the theoretical flux densities at these times provide a measure of the contribution from the eclipsed bright areas. Further consideration of this point will be deferred until the next section.

In Figure 7, the information given in Figure 6 is displayed in a manner that makes possible the isolation of the main component. In each section of Figure 7, recorded values of the relative power flux density at any phase of the eclipse are plotted against the flux density that would have been received at the same phase of the eclipse from a source on which the brightness was distributed in one of the six ways listed above.

If the brightness distribution on the sun were the same as that on any of the six postulated sources, then the resultant plot would be a straight line with unity slope. If a fraction  $x$  of the solar radiation originated in a manner corresponding to that from one of the postulated types of sources, while a fraction  $(1 - x)$  originated from small bright areas on the sun, then the curve to be expected would, at most points, have a slope  $x$  with sudden changes in slope corresponding to the effects of the eclipsing or uncovering of the small bright areas. The type of curve expected is shown in Figure 8.

An examination of the six sets of curves in Figure 7 shows that those in (a), (d), and (e) are inconsistent with the ideas expressed above and illustrated in Figure 8, and they must, therefore, be based on incorrect distributions of radio-brightness.



CHRISTIANSEN, YABSLEY, AND MILLS. - MEASUREMENTS OF SOLAR RADIATION AT  
A WAVELENGTH OF 50 CENTIMETRES DURING THE ECLIPSE OF NOVEMBER 1, 1948





Curves in (c) and (f), however, fulfil the requirement reasonably well in that they have a uniform slope over large sections and show a stepped approach to a straight line which has this slope and passes through the origin. At the maximum phase

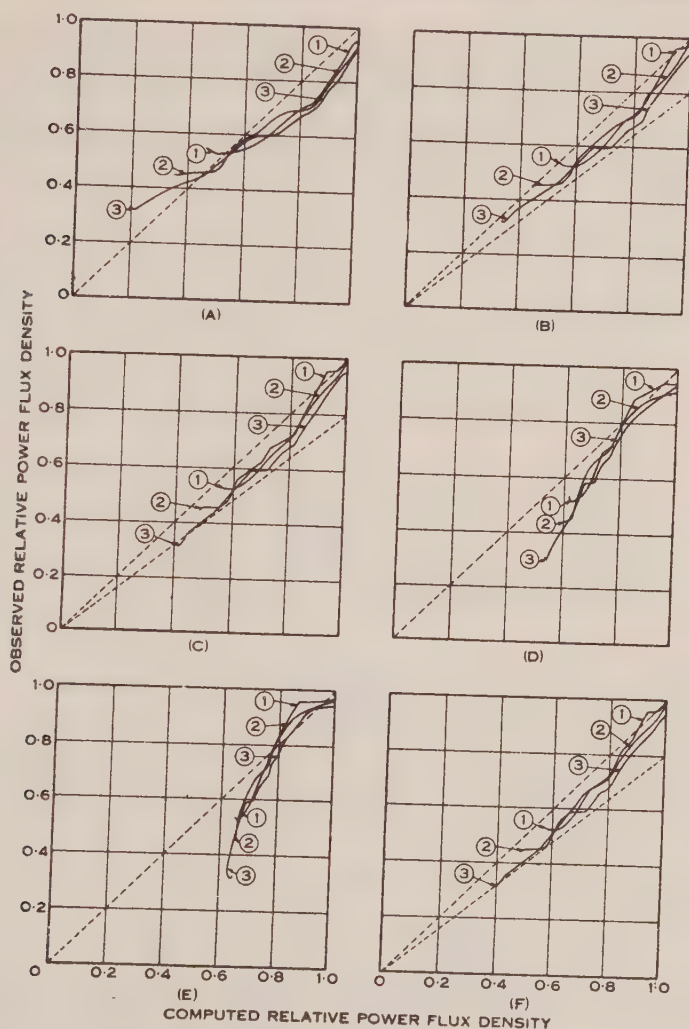


Fig. 7.—Measured values of power flux density during the eclipse, plotted against values calculated for the obscuring of various circularly-symmetrical sources, listed in Figure 6 and in the text. The labels 1, 2, and 3 refer to Sydney, Rockbank, and Strahan respectively.

of the eclipse at Strahan, when presumably all "bright" areas had been covered, the curves practically coincide with the straight line. The curves in (b) are considered to be less consistent with these postulates since, although there is some uniformity in general slope, the two latter features are not so well represented.

The straight line through the origin to which the curves in (c) and (f) tend has a slope of 0.8. This may be taken to represent the contribution of the symmetrically distributed component of the source: by implication the contribution of the localized bright patches was then 0.2.

Distributions of types (c) and (f) would produce very similar results during an eclipse, and it would be difficult to distinguish between them except near the point of first contact. In the three records shown here, the effects of a small bright area near the point of first contact are large enough to mask any smaller effects from the symmetrical source. Hence, no distinction can be made between the two postulated

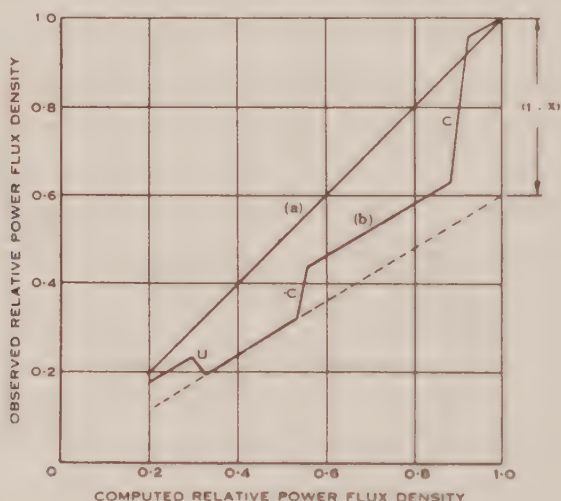


Fig. 8.—Ideal curves, of the type shown in Figure 7, when the assumed and actual general brightness distributions are the same.

(a) All the brightness is distributed in a circularly-symmetrical fashion.

(b) A fraction  $(1 - x)$  of the radiation originates in discrete bright areas.

C, Small area covered; U, Small area uncovered.

sources; neither can it be definitely concluded that "limb-brightening," predicted from theory, occurs. From the curves, it is possible, however, to draw the following conclusions:

1. On the day of the eclipse approximately 80 per cent. of the radio-frequency power received from the sun at a wavelength of 50 cm. originated in a manner consistent with either of two postulated distributions of radio-brightness, circularly symmetrical with respect to the centre of the disk. The first was a distribution derived from theoretical considerations and involved limb-brightening. The second was a uniform brightness distribution over a disk which had a radius 1.3 times that of the optical disk. In both cases about 40 per cent. of the received power originated outside the optical limb of the sun and it is to be expected that the observations

would be consistent with any symmetrical distribution which was approximately uniform across the optical disk and had 40 per cent. of its energy outside, but close to, the optical limb.

2. The other 20 per cent. of the received power originated in localized bright areas on and around the solar disk.

The curves marked "estimated for the 'quiet' sun" which are shown in Figures 2, 3, and 4 are for sources of type (f) contributing 80 per cent. of the total radiation.

It is worthy of note that the magnitude of the contribution of the small bright areas is consistent with that of the "variable component" (correlated with sunspot area) found by Pawsey and Yabsley(3). During the period of their observations the average value of the variable component of radiation accompanying a sunspot area of 85 parts in  $10^5$  of the visible solar disk was 17 per cent. of the total radiation. This is in fair agreement with the conclusion given above, that the contribution of the small bright areas was 20 per cent. of the total radiation.

#### (c) *The Size and Brightness of the Small Active Areas on the Solar Disk*

The conclusion that roughly 80 per cent. of the radio energy received from the sun had its origin in a symmetrical source provides a basis for estimating the size and brightness of the areas on the solar disk which were the sources of the residual 20 per cent. of the total radiation.

In the lower parts of Figures 2, 3, and 4 the slope curves for each of the experimental curves and also the slope curves relating to a symmetrical source of type (f) are shown. Assuming this source (or one of type (c), which has an almost identical slope curve) to be the origin of 80 per cent. of the solar radiation on the day of the eclipse, the covering and uncovering of small areas of more intense radio-brightness were represented on the experimental slope-time curve respectively by peaks above and dips below the theoretical curve. The width of each of the bright areas was calculated from the period of time over which the peak extended and the change in position of the moon during this interval. The widths so determined varied by little more than a factor of 2, the average value being 0.06 times the diameter of the optical disk. If each area is assumed to be circular, then the average bright area covered roughly 0.004 of the area of the visible disk.

The power radiated from each bright area was also calculated from the slope curves. The area bounded at the top by a given peak in the experimental slope curve and at the bottom by the theoretical slope curve is proportional to the energy in the bright region which produced the peak. By measuring each of the areas under the slope curve, and by taking account of the slope and time scales, the fraction of the energy from the whole of the sun, which originated in each bright region, was determined.

Eight bright areas were located, as shown in Figure 5. The values for the energy radiated from each, determined as above, ranged approximately from 0.03 to 0.05

of the total radiation at 50 cm., except for area 8 which contributed less than 0.01 of the whole. The average value was 0.035. Lehany and Yabsley(2) found that the temperature of the undisturbed sun (referred to the visible disk) was approximately  $0.5 \times 10^6$  °K. at a wavelength of 50 cm. If this value is used in conjunction with the magnitudes given above for the average radiated power and area of the active areas, it is found that the average effective temperature of these areas was approximately  $5 \times 10^6$  °K.

The effective temperature as determined individually for the active areas was found to vary by more than 10 to 1. It appears probable that only part of this variation can be attributed to experimental error, because the separate determinations for any one active area showed a maximum variation that was less than one-third of this. Hence it is reasonable to conclude that the active areas were not equally bright, and that the brightest had an effective temperature at least as high as  $10^7$  °K.

It may be noted that the total contribution of the small active areas, when determined by multiplying the average contribution per area by the number of discrete areas, has a value of 0.28, which is higher than the figure of 0.20 assumed earlier. This discrepancy arises from the fact that the slope of the experimental curves sometimes fell below that for the assumed symmetrical background source, even when no bright areas were being uncovered. This suggests that the contribution of the symmetrical background source may have been less than the value previously assumed, or that the background itself was "patchy."

#### (d) *Polarization Measurements*

Polarization changes during the first half of the eclipse are shown on the Melbourne record. Unfortunately recording of the two circularly-polarized components of the solar radiation was not possible at Sydney, and was effected for only a short period at Strahan.

Before the eclipse, the two components had amplitudes that differed by less than two parts in 100. During the course of the eclipse, relative changes in the two magnitudes were found. These changes were associated with the covering and uncovering of the small bright areas on the solar disk.

The changes were of short duration, the two components returning rapidly to equality. Apart from the short-duration changes, the two components of the solar radiation remained equal, within the range of experimental uncertainty of 3 per cent., during the obscuring of one hemisphere of the sun.

Smerd(11) has calculated that a central-dipole magnetic field on the sun, giving a field intensity of 50 gauss at the poles, would have caused the two components of the solar radiation (at 50 cm. wavelengths) to differ in magnitude by 20 per cent. at the maximum phase of the eclipse at Rockbank. A difference of 3 per cent. in the magnitudes of the two components would correspond to a field intensity of 8 gauss at the poles. It is concluded therefore that either the effect of the general magnetic



field was masked by the effects of the strong local field in the vicinity of the small active areas, or that the strength of the general magnetic field was considerably less than had been usually assumed.

The eclipsing of the active areas produced changes that sometimes were confined to one or other circularly-polarized component, or in some cases involved both components. An interesting feature of the polarization changes was that after the eclipse of an active area the predominance of one circularly-polarized component over the other did not persist for long; the magnitudes of the two components tended to approach equality.

## V. CONCLUSIONS

A study of the eclipse records has led to the following conclusions concerning the radio-frequency radiation from the sun, at a wavelength of 50 cm.:

1. On the day of the eclipse, there were present on or around the solar disk a number of small areas from which the intensity of radiation was considerably higher than that from the remainder of the solar disk. Such areas have been termed "bright" or "active" areas.

2. Seven of the eight or nine bright areas were found to be associated with optically visible features of the disk.

(a) Three were situated close to visible sunspots on the solar disk.

(b) Three were found to be close to positions on the solar disk occupied by sunspots 27 days earlier.

(c) One was found close to the position of a stable prominence on the southwest limb of the sun.

3. These bright areas contributed roughly one-fifth of the radiation at 50 cm. wavelength from the sun on the day of the eclipse.

4. An average bright area subtended a solid angle at the earth which was 0.004 of that of the visible solar disk.

5. The average effective temperature of such an area was about  $5 \times 10^6$  °K. The effective temperature of some of the bright areas probably exceeded  $10^7$  °K.

6. Four-fifths of the radiation from the sun, at the wavelength and time specified, appeared to originate in a source symmetrical with respect to the centre of the sun. The distribution of radio-brightness on this source was such that approximately 40 per cent. of the radiation from the sun originated outside the visible disk of the sun. The records were consistent with a theoretically-derived distribution involving limb-brightening, but the existence of limb-brightening was not proved.

7. The ratio of the contribution of the bright areas to that from the remainder

of the sun, when related to the sunspot area on the day of the eclipse, is in agreement with the statistical findings of Pawsey and Yabsley.

8. Relative changes in the magnitudes of the two circularly-polarized components of the radiation from the sun were found to occur as the "bright" or "active" areas on the sun were eclipsed. After such a change the magnitudes of the two components tended to return to equality.

9. There was no apparent tendency for the relative magnitudes of the two components to diverge as one hemisphere of the sun was covered. A difference of 3 per cent. between the two circularly-polarized components could have been detected at the maximum phase of the eclipse. A divergence greater than this might have been expected had there been a general magnetic field of the central-dipole type on the sun producing an intensity greater than 8 gauss at the poles.

## VI. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Radiophysics, C.S.I.R.O.

The authors wish to acknowledge the active participation of Mr. A. G. Little, Mr. A. J. Harragon, and Mr. J. D. Murray, in the preparations for the eclipse expeditions and in the observations during the eclipse. They wish also to acknowledge their indebtedness to the following: the staff of the Commonwealth Observatory, Mt. Stromlo, for information both on calculations of the circumstances of the eclipse and on optical observations of the sun during the eclipse; the staff of the Overseas Telecommunications Commission for assistance provided at the Beam Wireless Station at Rockbank; Mr. R. S. Millett, Mr. W. L. Clapham, and Mr. C. M. Gray for photographs; and Dr. J. L. Pawsey for valuable criticism of this paper during its preparation.

## VII. REFERENCES

- (1) REBER, G.—Solar radiation at 480 Mc/s. *Nature* **158**: 945 (1946).
- (2) LEHANY, F. J., and YABSLEY, D. E.—Solar radiation at 1200 Mc/s., 600 Mc/s., and 200 Mc/s. *Aust. J. Sci. Res. A* **2**: 48 (1949).
- (3) PAWSEY, J. L., and YABSLEY, D. E.—Solar radio-frequency radiation of thermal origin. *Ibid.* **2**: 198 (1949).
- (4) GINSBURG, V. L.—On solar radiation in the radio spectrum. *C. R. Acad. Sci. U.R.S.S.* **52**: 487-90 (1946).
- (5) MARTYN, D. F.—Temperature radiation from the quiet sun in the radio spectrum. *Nature* **158**: 632-3 (1946).
- (6) WALDMEIER, M., and MÜLLER, H.—Spectral energy distribution and centre-limb variation of radio frequency radiation from the corona. *Astr. Mitt. Zurich* **155** (1948).
- (7) SMERD, S. F.—Radio-frequency radiation from the quiet sun. *Aust. J. Sci. Res. A* **3** (in press).

- (8) COVINGTON, A. E.—Microwave solar noise observations during the partial eclipse of November 23, 1946.— *Nature* **159**: 405-6 (1947).
- (9) SANDER, K. F.—Radio noise from the sun at 3.2 cm. *Ibid.* **159**: 506 (1947).
- (10) ALLEN, C. W.—Variation of the sun's ultra-violet radiation as revealed by ionospheric and geomagnetic observations. *Terr. Magn. Atmos. Elect.* **51**: 1-18 (1946).
- (11) SMERD, S. F.—The polarization of thermal "solar noise" and a determination of the sun's general magnetic field. *Aust. J. Sci. Res. A* **3** (in press).

#### EXPLANATION OF PLATES 1-2

##### PLATE 1

Aerial used for reception of circularly-polarized waves.

##### PLATE 2

The sun on the day of the eclipse.

# SOLAR RADIATION AT A WAVELENGTH OF 10 CENTIMETRES INCLUDING ECLIPSE OBSERVATIONS

By J. H. PIDDINGTON\* and J. V. HINDMAN\*

(Plate 1)

[*Manuscript received August 3, 1949*]

## *Summary*

Observations were made of the intensity of solar radiation at a wavelength of 10.0 centimetres before and during the partial eclipse of November 1, 1948. Measurements were also made of the average level of solar radiation on a number of days before and after the eclipse. In this way conditions during the eclipse were related to the varying day-to-day level of intensity.

A determination of the distribution of intensity of radiation over the solar disk was made, the most intense radiation coming from near the limb and some radiation from beyond the limb. At least one small area was located which emitted more intensely than the average.

Polarization measurements, although inconclusive, suggest that radiation from one "hot" area was partially circularly polarized. Excess of either right- or left-hand polarized component at eclipse maximum was small, probably less than  $1\frac{1}{2}$  per cent. of the mean and perhaps zero. It is shown that the usually accepted value of the solar general magnetic field of 50 gauss at the poles should provide an excess of one component of about  $2\frac{1}{2}$  per cent. at eclipse maximum.

## I. INTRODUCTION

The theoretical distribution of the steady component of microwave radio-frequency radiation over the solar disk has been discussed by a number of workers(1-3) and shown to depend to a large extent on the temperature gradient in the chromosphere. The determination of such distributions at various radio frequencies is important because it provides data for the estimation of temperature gradients in the solar atmosphere. A solar eclipse provides an opportunity to make such a determination, and this paper is primarily concerned with a description of measurements of solar radiation at a wavelength of 10.0 cm. made in Sydney during the solar eclipse of November 1, 1948†. To relate conditions during the eclipse to those on other days, the intensity of solar radiation was measured on a number of days before and after the eclipse.

Previous measurements of solar radiation during an eclipse have been made by several observers including Covington(4) who used a wavelength of 10.7 cm. Covington's curve of intensity of radiation during the eclipse shows fluctuations of about 10 per cent. maximum intensity. These fluctuations may be caused by variation

\*Division of Radiophysics, C.S.I.R.O.

†Observations were also made at a wavelength of 50 cm. by Christiansen, Yabsley, and Mills and at 3.2 cm. by Minnett and Labrum of this laboratory. The results are in course of publication.



of intensity of radiation from the uneclipsed portion of the disk, or by other effects, but they render the results unsuitable for use in determining the distribution of the non-variable component of radiation over the solar disk. One of the principal objectives of the present investigation was to provide such a curve.

In addition to intensity measurements, the polarization of solar radiation was also examined at intervals during the eclipse, a method of excluding either right- or left-hand circularly polarized component being employed. It was hoped to associate certain polarization with given sunspots and also to measure the general magnetic field of the sun.

## II. APPARATUS

### (a) *The Radiometer and Associated Equipment*

The radiometer was similar to that already described(5) in connection with observations at 1.25 cm. wavelength except, of course, that all radio-frequency elements were designed for a wavelength of 10 cm.

The aerial unit is shown in Plate 1. It consists of a 68-in. paraboloid, *A*, mounted on a polar axis about which it may be driven either manually or by two gear boxes and an electric motor indicated as *B*. The aerial may be moved in declination by turning a lead screw, *C*.

Mounted on the paraboloid are the radio-frequency components, including the aerial feed horn, *D*, the chopping wheel driven by the motor, *E*, and the crystal mixer and the first stages of the intermediate-frequency amplifier in the box, *F*. Also shown in position is the screen, *G*, which allows right- or left-hand circularly polarized components of radiation to be received depending on its orientation. The telescope, *H*, is useful in quickly setting the aerial beam on the sun and in ensuring that the beam remains accurately set on the sun while the aerial is being driven.

The remainder of the radiometer has been described(5) including the calibrating devices which consist of two matched terminations replacing the aerial horn. One of these may be heated above ambient temperature, the rise being measured by a thermojunction and meter.

The aerial is mounted on a polar axis about which it is driven by a synchronous electric motor. In order to facilitate the initial setting of the aerial beam on the sun the gear ratio between the motor and the polar axis may be varied.

## III. EXPERIMENTAL METHOD

### (a) *The Absolute Intensity of Solar Radiation*

When the aerial beam of the radiometer was directed at the sun an increase of "aerial temperature" was noted. An absolute measurement of this change of aerial temperature,  $T_A$ , was made by using the hot and cold aerial terminations. This has been discussed previously(5), and also the fact that the equivalent black-body temperature of the sun,  $T_D$ , is related to  $T_A$  by the equation

$$T_D = FT_A,$$

where

$$F = 4\pi/G\Omega \text{ is the "aerial factor,"}$$

$G$  is the maximum gain of the aerial over an isotropic radiator, and  $\Omega$  is the solid angle subtended by the sun.

The aerial factor was determined by a novel method which may be used when the aerial is associated with a sensitive and stable radiometer. The method will be described in detail in a later paper dealing with equipment and techniques. Meanwhile, a very brief description is given here.

Let us imagine the aerial totally enclosed within an isothermal black body, so that the flux of energy received per unit solid angle is the same in every direction. We define two quantities  $A$  and  $B$  as follows:

Let  $A$  be the proportion of the total energy picked up by the aerial which is incident within the main lobe of the aerial. Obviously, then,  $(1 - A)$  is the proportion of energy picked up by the side and back lobes. Also, let  $B$  be the proportion of the energy received within the main lobe which is received within a cone of solid angle,  $\Omega$ , having its axis in the direction of maximum gain. It may easily be seen that

$$F = 1/AB.$$

The value of the factor  $A$  was found by measuring the increase in the amount of thermal radiation received (as indicated by the radiometer) when the main aerial lobe was directed first a little above the horizon, and then at the ground, just below the horizon. Virtually no radiation is received from the sky at this frequency and the thermal emission from the ground may be calculated when determinations of its temperature and emissivity have been made. The change in the amount of energy picked up in the side lobes as the main lobe is moved from sky to ground is small. A method of correcting for this effect will be given with the full description. It is obvious, then, that the increase in received radiation is proportional to  $A$  which may then be readily determined.

The value of  $B$  was found from a graphical analysis of two aerial response curves in two planes perpendicular to one another. These were obtained by using the sun as a point source radiator and allowing it to cross the main lobe of the aerial, first along a line of constant declination and then along a line of constant Right Ascension.

The value of  $F$  for the 68-in. disk was 133, the corresponding value of  $G$  being 1390. The value of  $F$  corresponds to a solar diameter of  $32'$ , all aerial temperatures being adjusted to correspond to this value.

#### (b) *Eclipse Observations*

To obtain a curve of the flux of solar energy during the eclipse, the aerial beam was driven so as to follow the sun from the commencement of the eclipse in Sydney (at 1647 hours, Eastern Australian Standard Time) until the sun was so close to the horizon that the record was no longer useful (1756 hours). Fortunately the sky was clear and following could be checked with the telescope. The intensity was recorded on the Esterline-Angus recorder at a paper speed of 12 inches per hour.

To check any drift in zero level, the beam was moved away from the sun for intervals of one minute each quarter-hour. It was moved  $15^\circ$  east of the sun's

position. The sensitivity of the equipment was measured before and after the eclipse. During the two-hour interval a change of 6 per cent. was noticed and this was assumed to have occurred at a uniform rate. Measurements of sensitivity on other occasions showed that changes usually occurred as slow drifts.

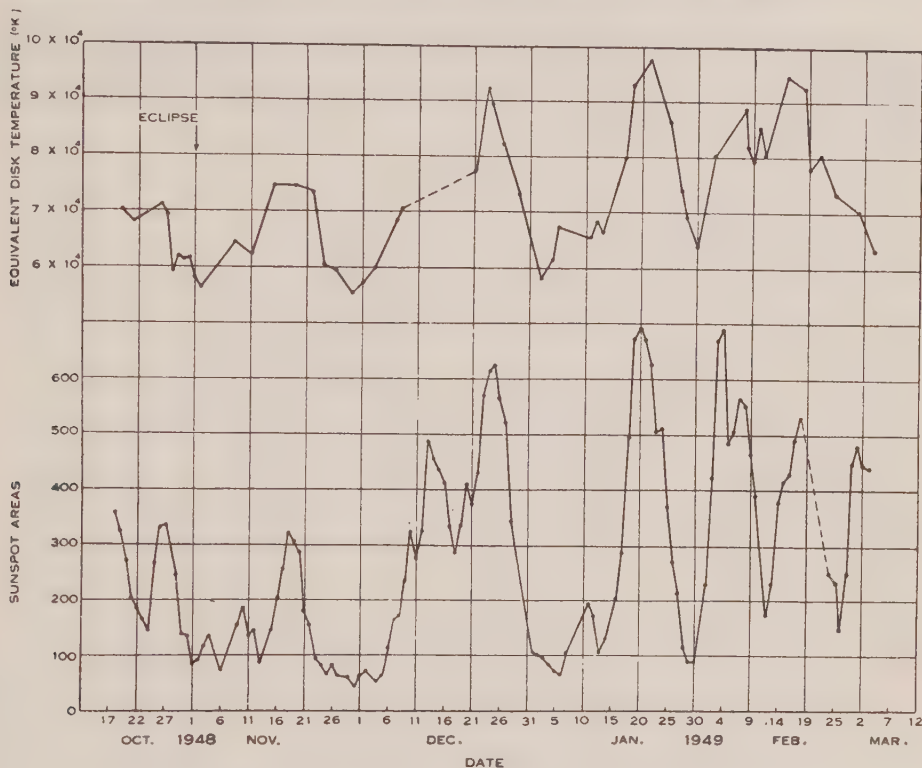


Fig. 1.—Solar radio-frequency radiation and sunspot areas.

Because of the presence of side lobes in addition to the main aerial beam it was necessary to make allowance for thermal energy radiated by the earth and picked up in these lobes while the main beam was directed above the horizon. The path followed by the sun during the eclipse was plotted and the main beam moved over this path at a period when the sun was in another part of the sky. The varying background radiation picked up in the side lobes was thus determined. A similar determination was made by moving the beam over a path displaced  $15^\circ$  east of the eclipse path. This displaced path corresponded to the directions of the aerial beam used for determining the background or zero level of radiation during the actual eclipse observations. The difference between the two curves of received radiation, thus determined, constituted the true correction to the eclipse intensity curve due to side lobe pickup.

Once just prior to the commencement of the eclipse, and three times during its progress, the polarizing screen was placed over the aerial and its position reversed four or five times. This allowed a determination of relative intensities of right-



and left-hand circularly polarized components.

#### IV. RESULTS

##### (a) Daily Measurements of the Level of Solar Radiation

Regular measurements of solar radiation were made from October 19, 1948, to March 3, 1949. One or more observations were made during a day and the average taken. The level was usually steady during a day but once or twice it changed by as much as 10 per cent.

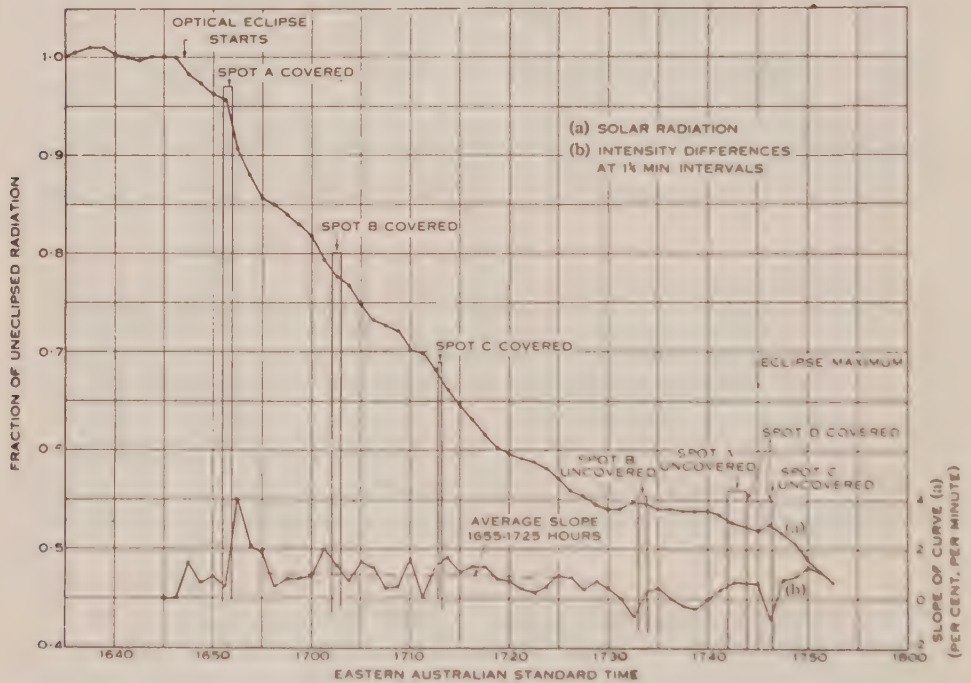


Fig. 2.—Eclipse observations.

(a) Solar radiation received during the eclipse period.

(b) Rate of change of radiation.

The intensity of radiation expressed as an average equivalent black-body temperature over the visible solar disk is plotted in Figure 1, sunspot areas for the same period being plotted for purposes of comparison. Similar observations at a wavelength of 10.7 cm. have been described by Covington(6), who found much larger variations, probably because his measurements were made during a period of more intense solar activity. Covington's minimum equivalent temperature was  $7.9 \times 10^4$  K. although an earlier figure(3) of  $5.6 \times 10^4$  K. was obtained when excess radiation from the vicinity of sunspots was not included.

If equivalent temperature is plotted against sunspot area for the period shown in Figure 1 a scatter of points is obtained among which a straight line may be drawn



so as to fit the points as well as possible. If this line is extrapolated to zero sunspot area we find the temperature to be  $5.4 \times 10^4$  °K. which is our estimate of the equivalent temperature of a completely "quiet" sun. This method is due to Pawsey and Yabsley(7) who used it to analyse Covington's results. They found a "quiet" sun temperature of  $6.5 \times 10^4$  °K. for a wavelength of 10.7 cm.

The highest temperature observed was  $9.8 \times 10^4$  °K., some 80 per cent. higher than the "quiet" sun value.

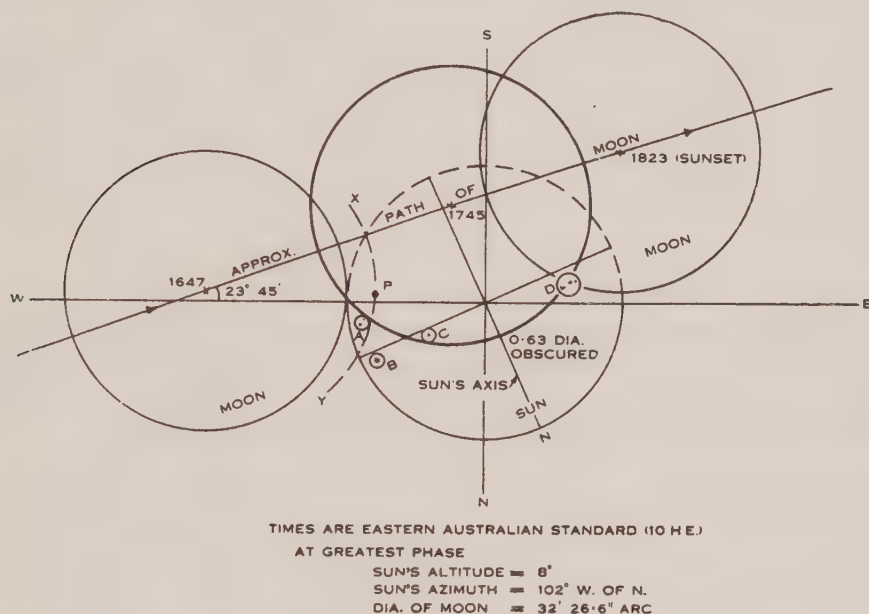


Fig. 3.—Circumstances of the eclipse.

The correlation between equivalent temperature and sunspot area is close, a fact which has been mentioned by Covington(6) and others.

The level of radiation at the time of the eclipse was about 8 per cent. above the minimum value, indicating that our intensity distribution results apply to a fairly quiet sun.

(b) *Intensity Measurements during the Eclipse*

The sun's radiation intensity over a period from a few minutes before visual contact until past maximum are shown in Figure 2 curve (a). Intensities are given in arbitrary units and are plotted as  $1\frac{1}{2}$  minute averages; they are corrected for change in receiver sensitivity, zero drift, and change in background temperature.

Circumstances of the eclipse are shown in Figure 3, all times being Eastern Australian Standard. It will be observed that the optical eclipse commenced at 1647 hours and maximum was reached at 1745 hours. Shortly after this time the sun was so close to the horizon that no further useful results were obtainable.

In plotting curve (a) of Figure 2, intervals of  $1\frac{1}{4}$  minutes were used so that the small irregular fluctuations in receiver output were smoothed and the work associated

with tabulating the results and applying the various corrections mentioned in Section III (b) was limited. However, the original data were in the form of a continuous curve and the form of this curve for the few minutes around the time of commencement of the eclipse is similar to that of curve (a) of Figure 2, the irregular fluctuations being about  $\pm 0.007$  on the scale of this figure. It would appear, therefore, that the radio eclipse did not commence before the visual eclipse by more than about 30 seconds. The period of 30 seconds is the best, although somewhat doubtful, estimate of the interval between the two commencements.

This result differs from that of Covington(4) who concluded that the radio eclipse commenced three minutes before the optical eclipse. Covington's observations were made on a much more active sun, however, and could be explained as due either to radiation from a high level or to fluctuations of total solar radiation.

It may also be observed that small fluctuations in intensity of radiation occurred some eight minutes before commencement of the optical eclipse. For several minutes before, however, the curve is quite level.

### (c) Polarization Measurements

Polarization measurements were made at the times indicated in Table 1 on the day of the eclipse and on several days following. The quantity measured which we call  $P$  is the excess of the left-hand (LH) or right-hand (RH) circularly-polarized component\* expressed as a percentage of the total received radiation.

TABLE 1

Time, E.A.S.T.	P-Nov. 1		P-Nov. 2		P-Nov. 3
1630	LH	1	RH	1	RH 2
1704	All	0	LH	1	RH 1
1730	LH	2	LH	2	LH 1
1742	LH	1	LH	2	LH 1
1747	LH	3	LH	1	LH 1

The accuracy of these measurements was high so that a small rapid change of polarization could be measured with an accuracy of better than one per cent. However, over long periods, a drift in absolute level of polarization often occurred so that day-to-day values varied by several per cent. This was due to a combination of two factors: a real change of polarization of solar radiation and instrumental effects. This slow drift was not important in interpreting the eclipse results.

The values of  $P$  given in Table 1 are plotted in Figure 4, which will be discussed later.

## V. INTERPRETATION OF RESULTS

### (a) Sunspot Areas and Areas of High Emission

During the eclipse in 1946, Covington(4) associated the disappearance of a large group of spots with a drop in flux of 25 per cent. and estimated the equivalent

\*The polarization is right- or left-handed according to the direction of rotation of the electric vector at a fixed plane at right angles to the direction of propagation when viewed along the direction of propagation.

temperature of the highly emissive or "hot" area to be about  $1.5 \times 10^6$  °K. During the eclipse in 1948, however, the few visible sunspots were small and, judging by the curve of equivalent disk temperatures shown in Figure 1, not very effective in increasing radiation. This is borne out by the fact that no clearly marked changes in slope of the curve (a) of Figure 2 may be associated with the appearance or disappearance of sunspots.

However, an inspection of curve (b) of Figure 2 which shows the rate of decrease of solar flux received during the eclipse period indicates that a "hot" area was obscured between about 1651 and 1700 hours. At first sight this might be identified with the sunspot group A which, as shown in Figure 2, was covered at about 1651-1652 hours. There are several reasons, however, for believing that this is not the case.

The obscuration of the "hot" area commenced when the limb of the moon was in the position XY of Figure 3. All points along this arc are covered at the same instant but are uncovered at different times. If the "hot" area were near spot A it would be uncovered commencing at about 1743 hours. There is no sign of a sudden increase in radiation at that time. Observing from three widely separated sites at a frequency of 600 Mc/s., Christiansen, Yabsley, and Mills(8) were able definitely to locate a number of "hot" areas observed during the November eclipse. One of these is indicated by the point P on Figure 3 and there is little doubt that this is the "hot" area discussed above, particularly since Christiansen, Yabsley, and Mills found no radiation from spot A. This area, P, was not uncovered until about 1755 hours and so the effect was not observed on our record.

Some of the other variations in amplitude of curve (b) of Figure 2 are probably due to the covering or uncovering of "hot" areas since the amplitude of such variations is rather larger than usually obtained when observing the uneclipsed sun.

If the sudden increase in slope of curve (a) of Figure 2 at 1652 hours is due to the obscuration of a "hot" area then from the period taken to cover the area and the excess change in received flux an estimate may be made of the apparent temperature of the area. If the area is circular then its area will be about  $2 \times 10^{-3}$  area of the solar disk and its temperature about  $10^6$  °K. If the small discontinuities at 1701 and 1743 hours are due to "hot" areas being obscured, then the corresponding temperatures would be about  $0.2 \times 10^6$  °K. and  $0.5 \times 10^6$  °K. respectively.

### (b) Polarization of "Hot" Areas

Owing to the apparently random variations of polarization of the order one per cent. and the rather long intervals between measurements, it is impossible to attribute any definite degree of polarization to any particular area of the solar disk. An inspection of Figure 4 suggests two things however: continuous measurement of polarization was in progress between 1743 and 1747 hours and the change of polarization by 2 per cent. took place within one minute. This change took the form of an increase in LH polarization while the RH component remained unchanged thus indicating that the change was probably due to the uncovering of an area with predominantly LH polarized radiation. It appears likely from an inspection of Figure 4 that the area providing the LH polarized radiation was eclipsed before 1704 hours. It could not have been the area P, however, as this was not uncovered until some minutes after the observed change in polarization.



*(c) Distribution of Background Radiation over the Disk*

Curve (a) of Figure 2 might be compared with various theoretical eclipse curves corresponding to different distributions of radiation over the solar disk. However, since it has reasonably been established that there is one and probably more small localized "hot" areas, it was considered desirable to remove the effects of these before making the comparison.

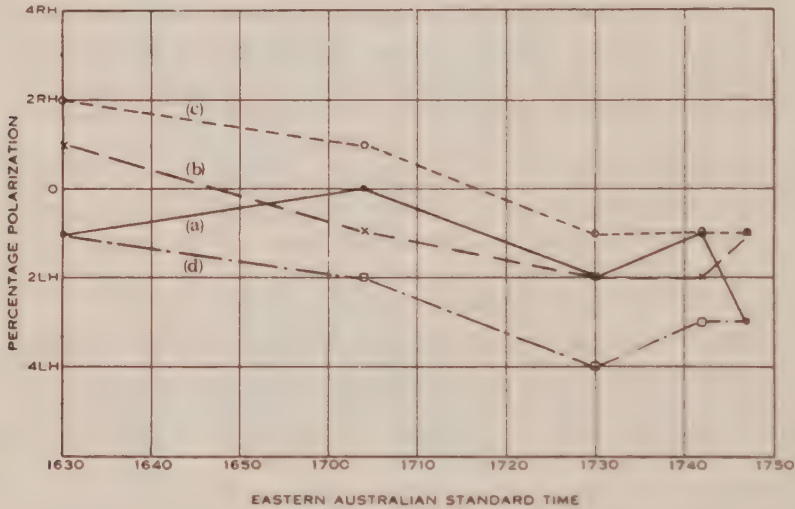


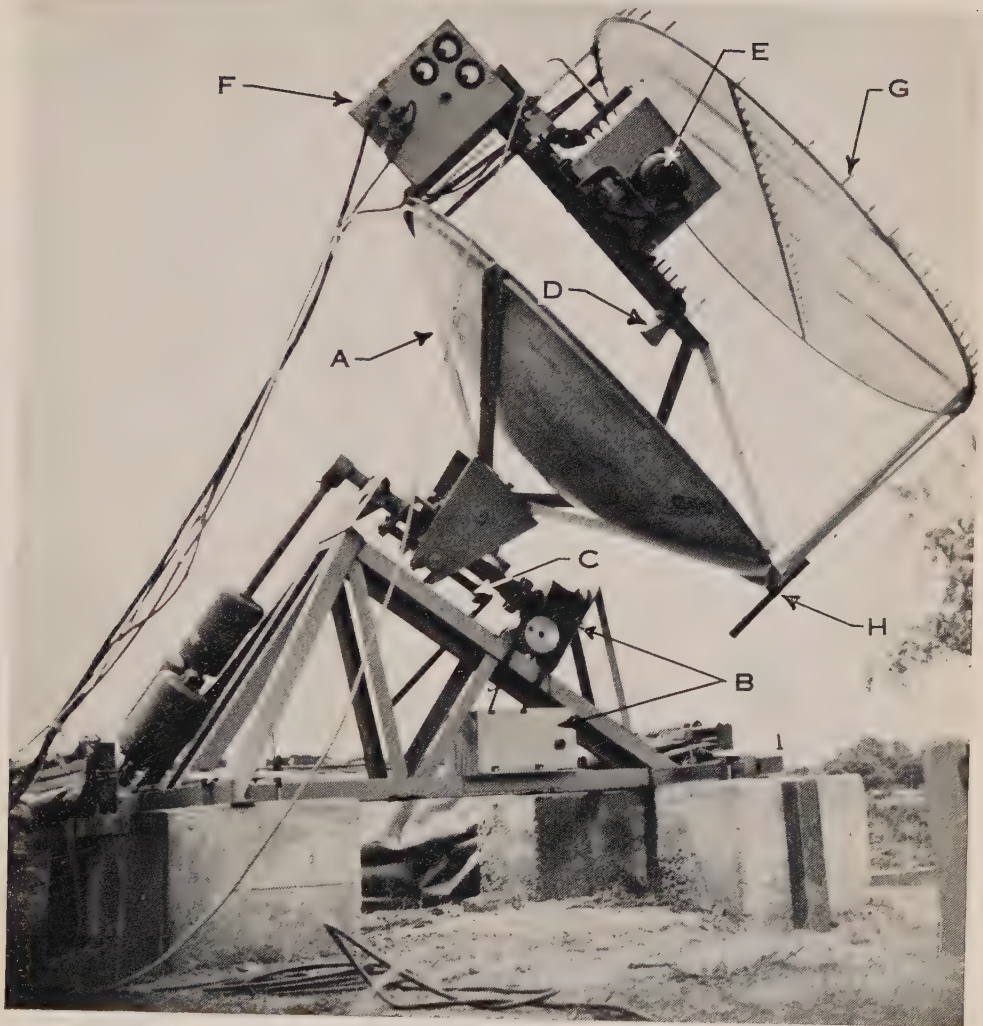
Fig. 4.—Variations in polarization.

- (a) During November 1.
- (b) During November 2.
- (c) During November 3.

The levels of radiation before the start of the eclipse and at the maximum were retained, the latter being an averaged value over the five minute period centred at eclipse maximum. Sudden changes of slope of the curve such as at 1652 hours smoothed out so that the level of curve (b) of Figure 2 was kept constant over short periods.

The new eclipse curve is shown as curve (a) of Figure 5. Also shown in Figure 5 are theoretical curves corresponding to uniform distributions over the disk from which radio-frequency radiation is received (curve b) and over a narrow ring around the circumference of this disk (curve c). The experimental curve (a) is obtained, approximately, if a combination is made of 32 per cent. of curve (c) and 68 per cent. of curve (b). A distribution of radiation consisting of 32 per cent. from a thin ring around the solar limb and 68 per cent. from a uniform disk will, therefore, provide an eclipse curve very similar to that observed. A departure from the composite ring and disk distribution may be observed in the early stages of the eclipse where the necessary proportion of ring to give the observed curve is only 19 per cent. at 1650 hours. This result may be explained if the radiation which, in order to simplify analysis is assumed to originate in a thin ring, is really distributed over a finite area near the solar limb.





PIDDINGTON AND HINDMAN. — SOLAR RADIATION AT A WAVELENGTH OF 10 CENTIMETRES  
INCLUDING ECLIPSE OBSERVATIONS



From the relative slopes of the observed eclipse curve and that due to a uniform disk it is possible to estimate, approximately, the equivalent temperature at the limb of the sun. One minute after commencement of the radio eclipse the observed drop in radiation is about ten times that for a uniform disk. We have seen that a uniform disk (the radio and optical disks are so nearly equal that a correction is not necessary) would have a temperature of about  $6 \times 10^4$  °K. so that the limb temperature, averaged over the area obscured by the moon in the first minute, is  $6 \times 10^5$  °K. The average for the first two minutes is about  $2 \times 10^5$  °K.

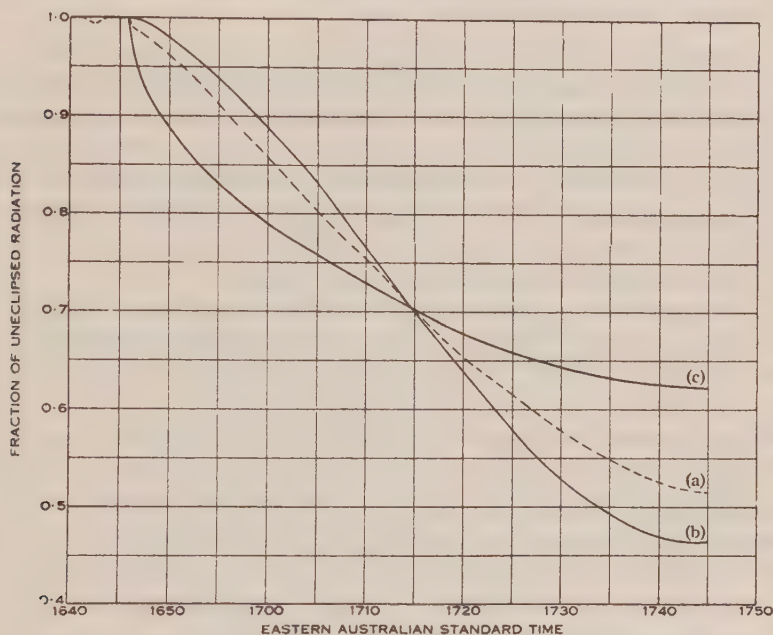


Fig. 5.—Eclipse curves.

- (a) The smoothed intensity curve.
- (b) The theoretical curve for a uniform disk.
- (c) The theoretical curve for a circumferential ring.

We have seen in Section IV (b) that the best estimate of time of commencement of the radio eclipse is 30 seconds before the visual eclipse. The rate of movement of the moon along the solar radius at the point of contact is about  $2 \times 10^4$  km. min.<sup>-1</sup>, which suggests that the highest level from which intense radiation is received is about 10,000 km.

The effect of diffraction of radiation at the edge of the lunar disk has been checked and found to correspond to a displacement of about 2000 km. on the solar disk. This is not important in connection with either the highest level of radiation or the position of radiation associated with spots.

#### (d) *The Sun's General Magnetic Field*

Radio-frequency energy emerging from an ionized atmosphere in the presence of a magnetic field may have two components, one right- and one left-hand circularly

polarized. These two components will have different absorption coefficients and hence different average levels of origin, the more strongly absorbed ray emerging from a higher average level. At the level of origin of 10-cm. radiation in the solar atmosphere, a positive upward temperature gradient will result in greater emission of the magneto-ionic component which has the larger absorption coefficient. A quantitative treatment given in Appendix I shows that if the sun has a general magnetic field of strength 50 gauss at the poles and the southern half of the disk is obscured, the remaining radiation should have a right-hand component about 2.5 per cent. stronger than the left-hand.\*

An examination of Figure 4 for change of polarization between 1704 and 1742 hours on November 1 during which period most of the change should have occurred is inconclusive. The polarization changed to LH by one per cent. while on the following two days it changed by an average of 1.5 per cent. The reason for the changes when an eclipse was not in progress was instrumental, caused by an increasing proportion of aerial side-lobes being directed onto the ground as the zenith distance of the sun increased. It is worth noting that if, as discussed above, an area emitting 2 per cent. excess LH component were eclipsed on November 1 between 1704 and 1745 hours then in the absence of this effect the eclipse polarization curve would have been curve (*d*) of Figure 4. The change in general level of the curve is not important; such an effect has been frequently observed and associated with movement of sunspots across the disk. The general form of the curve is very similar to that for the following two days and suggests that the change of polarization due to the eclipse was not as great as 2 per cent. and may have been zero.

We conclude that, although the results are not definite, they suggest that if a general magnetic field exists at all it is considerably smaller than the usually accepted value of 50 gauss at the poles.

This result only applies, of course, to the region from which 3000 Mc. s. radiation emerges. Such region, according to recent calculations(3), should have a mean electron temperature and density of about  $5.5 \times 10^4$  K. and  $1.5 \times 10^9$  cm.<sup>-3</sup> respectively and be situated at a level of about 11,000 km. above the photosphere.

It has often been suggested that the general magnetic field of the sun, while having a value of about 50 gauss at the photosphere is radially limited so that its value is small at levels of a few thousand kilometres. Cowling(9) has discussed some early theories and Kiepenheuer(10) has more recently considered some aspects of limitation by electric currents.

## VI. CONCLUSIONS

(i) During the period October 19, 1948, to February 4, 1949, the average equivalent black-body temperature of the solar disk varied from  $5.6 \times 10^4$  K. to  $9.8 \times 10^4$  K., having a value of  $5.8 \times 10^4$  K. during the eclipse of November 1, 1948.

(ii) The distribution of radiation over the disk at the time of eclipse (a "quiet" period) could be represented by a model sun having 68 per cent. of its radiation

\*Provided the sun's north pole is a north seeking pole, that is, a south magnetic pole.



emitted uniformly over the visible disk and 32 per cent. concentrated near the limb.

(iii) At least one small area was found which emitted excess radiation, having an equivalent temperature of the order  $10^6$  °K. The results also suggested that partially circularly-polarized radiation was emitted by another small area.

(iv) An attempt was made to measure the general magnetic field of the sun and the results, although inconclusive, indicated that if such a field existed at all it had a value considerably smaller than the usually accepted value of 50 gauss at the poles.

## VII. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Radiophysics, C.S.I.R.O. The authors wish to thank Dr. J. L. Pawsey for useful discussions. Much of the equipment used in the investigation was developed and tested by Mr. H. C. Minnett and Mr. N. R. Labrum, and Mr. G. H. Trensky assisted in making the observations.

## VIII. REFERENCES

- (1) MARTYN, D. F.—*Proc. Roy. Soc. A* **193** : 44 (1948).
- (2) WALDMEIER, M., and MÜLLER, H.—*Astr. Mitt. Zurich* No. 155 (1948).
- (3) PIDDINGTON, J. H.—*Proc. Roy. Soc. A* (communicated).
- (4) COVINGTON, A. E.—*Nature* **159** : 405 (1947).
- (5) PIDDINGTON, J. H., and MINNETT, H. C.—*Aust. J. Sci. Res. A* **2** : 63 (1949).
- (6) COVINGTON, A. E.—*Proc. Inst. Radio Engrs.* **36** : 454 (1948).
- (7) PAWSEY, J. L., and YABSLEY, D. E.—*Aust. J. Sci. Res. A* **2** : 198 (1949).
- (8) CHRISTIANSEN, W. N., YABSLEY, D. E., and MILLS, B. Y.—*Nature* (in press).
- (9) COWLING, T. G.—*Mon. Not. R. Astr. Soc.* **92** : 407 (1932).
- (10) KIEPENHEUER, K. O.—*Ann. Astrophys. Paris* **9** : 42 (1946).
- (11) BOOKER, H. G.—*Proc. Roy. Soc. A* **147** : 352 (1934).
- (12) BABCOCK, H. W.—*Astrophys. J.* **105** : 105 (1947).

## EXPLANATION OF PLATE I

The radiometer aerial and polarizing screen.

## APPENDIX I

Using the notation of Booker(11) the propagation of a plane monochromatic electromagnetic wave through a homogeneous ionized medium under the influence of a superimposed magnetic field is described by the complex refractive index

$$q^2 = 1 - \frac{2x}{2(1 - iz) - \frac{y_T^2}{1 - x - iz} \mp \left[ \frac{y_T^4}{(1 - x - iz)^2} + 4y_L^2 \right]^{\frac{1}{2}}}$$

The *power\** absorption coefficients  $\kappa_o$  and  $\kappa_e$  for the ordinary and extraordinary ray components are obtained by substituting  $q = \mu - ic\kappa/2p$  in the above equation and solving for the two values of  $\mu$  and  $\kappa$ .

In general the values of  $\kappa_o$  and  $\kappa_e$  differ so that the radiation corresponding

\*Twice the value given by Booker who deals with ratios of electric field strengths.

to the ordinary and extraordinary rays might be expected to experience different degrees of absorption (and hence by Kirchhoff's law, of emission) at a given level in the solar atmosphere. Thus, as shown by Martyn(1), the intensity of the two rays emerging from the solar atmosphere will differ.

At the levels from which appreciable radiation of wavelength 10 cm. emerges it has been shown(3) that the values of  $x$ ,  $y$ , and  $z$  in the above equation are small. Under these conditions we may estimate the small difference in  $\kappa$  for the two rays as a proportion of  $\kappa_0$ . If we put  $(\kappa_e - \kappa_o) = \delta\kappa$ , then it may be shown that\*

$$\frac{\delta\kappa}{\kappa} = 4p_L/p \quad \dots\dots\dots(1)$$

In those regions where  $x$ ,  $y$ , and  $z$  are small the absorption coefficient is given approximately by the Lorentz formula

$$\kappa = \frac{4\pi N e^2 \nu}{m c p^2}, \quad \dots\dots\dots(2)$$

this value applying with sufficient accuracy for the ordinary and extraordinary rays and for any combination of the two.

It is evident that a small change in angular frequency,  $\delta p$ , in equation (2) will cause a small change in absorption coefficient. Thus

$$\frac{\delta\kappa}{\kappa} = - \frac{2\delta p}{p} \quad \dots\dots\dots(3)$$

We have seen by equation (1) that the absorption coefficients of the ordinary and extraordinary rays differ by a small amount. We may, therefore, by combining equations (1) and (3) express the difference in absorption of the two rays in terms of a small difference in frequency:

$$\begin{aligned} \frac{\delta p}{p} &= - 2p_L/p \\ &= - \frac{2He \cos \theta}{mcp}, \quad \dots\dots\dots(4) \end{aligned}$$

where  $\theta$  is the angle between the direction of propagation and that of the imposed magnetic field  $H$ .

Having expressed the difference in properties of the ordinary and extraordinary rays (as far as absorption is concerned) in terms of an equivalent small difference in frequency it is now necessary to find the relationship between the equivalent temperature of a given area of the solar disk and the frequency. In this way we find the small change in equivalent temperature of each element of the solar disk caused by a small change in frequency, and so from (4) we have the difference in equivalent temperature for the ordinary and extraordinary rays.

It has been shown(3) that the equivalent black-body temperature of the solar

\* This calculation was made by Mr. K. C. Westfold of this Laboratory.

disk at radio frequency  $p/2\pi$  is given approximately by the empirical formula

$$T(\varphi) = 4800 + \frac{\alpha}{p} \sec \frac{1}{2}\varphi + \frac{\beta}{p^2} \sec \varphi, \quad \dots\dots\dots(5)$$

where  $\alpha$  and  $\beta$  are constants given and  $\varphi$  is the angle between the direction of emission of radiation and the solar radius at the place of origin of the radiation. By integration we find the average equivalent temperature over the disk to be

$$T_D = 4800 + \frac{4\alpha}{3p} + \frac{2\beta}{p^2} \quad \dots\dots\dots(6)$$

In order to simplify the treatment we will separate each of the last two terms of equation (6) into two components, one corresponding to uniform equivalent temperature over the disk and the other to a peripheral concentration of high equivalent temperature. Since  $H \cos \theta$  is zero at the limb of the sun (if an axial dipole field is assumed) the limb component of radiation will be equal for the ordinary and extraordinary rays so that this component may be neglected and radiation assumed to originate in a uniform disk. It may be shown that of the term  $4\alpha/3p$  of equation (6), 75 per cent. corresponds to uniform radiation and 25 per cent. to peripheral radiation. The term  $2\beta/p^2$  is separable into 63 per cent. uniform distribution and 37 per cent. peripheral radiation. Thus we may write

$$T_D = T_1 + T_2 + T_3, \quad \dots\dots\dots(7)$$

where  $T_1 = 0.75 \times 4\alpha/3p$ ,  $T_2 = 0.63 \times 2\beta/p^2$ , and  $T_3$  is the sum of all radiation which is either independent of the frequency or originates near the solar limb. Now a small change in angular frequency will cause corresponding changes in  $T_1$  and  $T_2$  such that

$$\frac{\delta T_1}{T_1} = -\frac{\delta p}{p} \text{ and } \frac{\delta T_2}{T_2} = -\frac{2\delta p}{p} \quad \dots\dots\dots(8)$$

In considering small equivalent temperature differences between the right- and left-hand circularly-polarized components, the term  $T_3$  will play no part since propagation is either at right angles to the magnetic field or is independent of frequency. Thus we may write

$$\begin{aligned} \frac{\delta T_D}{T_D} &= \frac{\delta T_1 + \delta T_2}{T_D} \\ &= -\frac{\delta p}{p} \cdot \frac{T_1 + 2T_2}{T_D}. \end{aligned}$$

Substituting(3) values for  $\alpha$ ,  $\beta$ , and  $p$  we find  $T_1 = 30,000^\circ\text{K.}$ ,  $T_2 = 6000^\circ\text{K.}$ , and  $T_D = 54,000^\circ\text{K.}$ , so that

$$\delta T_D/T_D = -0.78 \delta p/p.$$

Hence, from equation (4), the small difference in equivalent temperature of the solar disk for the two magneto-ionic components is given by

$$\frac{\delta T_D}{T_D} = \frac{1.56 H e \cos \theta}{mcp} \quad \dots\dots\dots(9)$$

An estimate of the average value of  $H \cos \theta$  over the northern half of the solar disk may be made by assuming the magnetic field to be due to a dipole at the centre providing a field of 50 gauss at the poles (see, for example, Babcock 12). It is found that the average value of  $H \cos \theta$  is 18.7 gauss. Substituting in equation (9) we find

$$\delta T_D / T_D = 0.027.$$

At the frequency concerned (3000 Mc/s.) and for a "quiet" sun the general magnetic field at the poles is given by

$$H = 1850 \delta T_D / T_D.$$

There is one factor of which we have not taken account, that is, the area eclipsed is not half a circular disk as assumed but has a shape as shown in Figure 3. Since the difference in the areas concerned is not large and the value of  $H \cos \theta$  is small near the equator, the correction is small enough to be neglected.

By reference to the equation indicating the sense of polarization of the two magnetoionic components (Booker 11), it may be shown that the left-hand component is more absorbed when the directions of propagation and the magnetic field coincide. At eclipse maximum the radiation leaves the northern hemisphere which is thought to be a south magnetic pole. The right-hand component should, therefore, be more absorbed and hence provide a higher equivalent temperature.



# SOLAR RADIATION OF WAVELENGTH 1.25 CENTIMETRES

By J. H. PIDDINGTON\* and H. C. MINNETT\*

[*Manuscript received April 22, 1949*]

## *Summary*

Observations of solar radiation at a wavelength of 1.25 cm. and covering a period of about six months are described. The average intensity of radiation corresponded to a black-body temperature of  $1.00 \times 10^4$  °K. with a maximum error of about  $\pm 5$  per cent.

Day-to-day variations in intensity were less than  $\pm 3$  per cent., which was the limit of observational accuracy. Short-period fluctuations were observed on a few occasions; even during periods of intense solar activity they were not greater than  $\pm 5$  per cent.

The distribution of intensity over the solar disk was measured by a method analogous to the Michelson interferometer technique and found to be consistent with 84 per cent. of the radiation coming from a uniform disk and 16 per cent. from a narrow ring around the circumference.

## I. INTRODUCTION

Since radio-frequency radiation from the sun was first observed by Southworth and Reber(1, 2) numerous investigators have measured solar radiation at wavelengths ranging from about 1.2 centimetres to many metres. That the emission of such radiation is a complex phenomenon and cannot be entirely due to simple thermal processes was first suggested by Appleton(3). Using a wavelength of 1.5 m. Pawsey, Payne-Scott, and McCready(4) showed that there is a correlation between radiation and the presence of sunspots. Subsequently Pawsey, Payne-Scott, and McCready(5) and also Ryle and Vonberg(6), using a wavelength of 1.7 m. demonstrated that a large proportion of solar radiation may originate in the vicinity of sunspots.

The majority of observations of solar radio radiation have been made at wavelengths greater than 50 cm., the source of most of the radiation being in the corona(7, 8). At microwavelengths much of the radiation originates in the chromosphere and as some of the outstanding problems of solar physics are associated with this region, observations at these wavelengths should be particularly useful.

### *(a) The Components of Radiation*

At microwavelengths(9) solar radiation is found to consist of a more or less steady component which may vary slowly over a period of days and a second component which appears only rarely and has a duration of a few minutes.

The shortest wavelength for which a reasonably extensive series of observations has been published is 10.7 cm. At this wavelength Covington(9) measured average daily levels of radiation for several months. These levels expressed in terms of

\*Division of Radiophysics, C.S.I.R.O.

equivalent black-body radiation had a minimum of about  $7.9 \times 10^4$  °K. and often rose above twice this value. Frequently the level was fairly steady for several days in succession and showed a marked correlation with relative sunspot numbers. In addition to the slowly varying general level of radiation, Covington occasionally observed "bursts" lasting a few minutes and raising the equivalent black-body temperature of the sun by amounts varying from a few thousand degrees to about  $1.7 \times 10^5$  °K.\* The times of occurrence of these bursts correlated with sudden ionospheric disturbances.

Solar radiation at a wavelength of 1.25 cm. has been measured on two occasions by Dicke and Beringer(10), levels corresponding to black-body temperatures of  $1.0 \times 10^4$  and  $1.1 \times 10^4$  °K. being found.

Our observations were made at a wavelength of 1.25 cm. They consisted of a determination of the steady level each day and a series of recordings of the radiation intensity for periods of several hours on some days. Efforts were made to correlate these observations with other solar phenomena.

### *(b) The Distribution of Radiation over the Disk*

It has been shown by Ginsburg(7) and Martyn(8) that in the absence of disturbances in the solar atmosphere radio-frequency energy received from the sun is probably thermal in origin and originates at well-defined levels in the solar atmosphere. At microwavelengths the solar atmosphere is a black body and radiation emerges from a range of levels appropriate to the frequency and angle of emergence. Since the temperature at a given level does not vary rapidly the amount of radiation from the "quiet" sun at any given frequency tends to remain fairly constant.

At optical wavelengths the sun's visible disk is brightest at the centre and darkest at the limb, a consequence of the negative outward temperature gradient at levels from which visible radiation emerges. Because of the positive outward temperature gradient at the levels from which microwave radio energy emerges, such radiation should, as shown by Martyn(8), be most intense from the limb of the sun and least intense from the centre.

Observations of radiation during a partial eclipse have been made by Dicke and Beringer(10), whose results approximately fit the assumption of a uniformly radiating disk. The points on their curve are rather scattered, however, and do not permit an accurate quantitative estimate of distribution of intensity of radiation.

Using an interferometer technique at a wavelength of 1.25 cm. the intensity distribution is investigated experimentally and the results are shown to indicate that radiation from near the limb is more intense than radiation from near the centre of the disk.

## II. EXPERIMENTAL METHOD

The apparatus and observing technique were similar to those briefly described in an earlier paper(11) dealing with microwave radiation from the moon. A more

\*One burst of  $3 \times 10^6$  °K. is reported, but this may be a misprint.

detailed description of microwave radio-astronomy techniques is in course of preparation.

(a) *The Microwave Temperature of the Disk*

The method of measuring the temperature of the solar disk was similar to that used when observing the moon. Receiver sensitivity, radiation received, and atmospheric absorption were measured and permitted the calculation of an equivalent aerial temperature. Knowledge of the width of aerial beam and solid angle subtended by the sun allowed an estimate to be made of the average temperature over the solar disk at a wavelength of 1.25 cm., assuming it to be a black body. This is called the microwave temperature of the disk.

(b) *Continuous Recording of Radiation Intensity*

When a continuous record of solar radiation was required a fan-shaped aerial beam was used and the aerial driven about a polar axis by a synchronous motor. The beam was several degrees wide in the direction of changing hour angle so that errors in the speed of driving owing to back-lash and varying electric mains frequency were not important. The beam was turned away from the sun at half-hour intervals in order to check the zero reading. Measurements of receiver sensitivity and atmospheric absorption were made at intervals of a few hours.

(c) *Distribution of Radiation over the Disk*

The method used for determining the distribution of radiation over the disk was essentially Michelson's interferometry method of measuring stellar diameters. The method has been used to locate areas of intense solar activity(6) and its extension to the determination of the general distribution has been suggested(5).

In practice the method consists of receiving solar radiation at two aerials spaced a large number of wavelengths apart. The composite aerial diagram is then a number of narrow fan-shaped lobes whose separation may be varied by changing the aerial separation.

Two 18-in. paraboloids were used, spaced 53 in. apart so as to form an interference pattern with minima a little more than a solar diameter apart. The main lobe of the aerial diagram of each paraboloid was directed at the sun and the gain of the two aerials was adjusted to equality by an attenuator in one of the waveguide feeds. As the sun moved across the composite aerial diagram the oscillations in the receiver output were plotted by a recording milliammeter.

By moving the aerial beam so that an interference pattern was produced by a signal generator placed a few hundred yards away, it was possible to check that the signal dropped substantially to zero at the minima when a point source was observed. In addition, the positions of the various maxima and minima were recorded and related to the corresponding maxima and minima caused by the sun in passing through the aerial beam. The amplitude of the interference oscillation depended on the distribution of radiation over the solar disk and could be compared with theoretical curves derived by making various assumptions about such distribution.



## III. RESULTS AND THEIR INTERPRETATION

The results consist of spot measurements of solar radiation, usually made near midday, between April and August 1948. In addition, continuous recordings of solar radiation were made for several hours on a number of days between July and October. Separate experiments were carried out to determine the distribution of intensity over the disk.

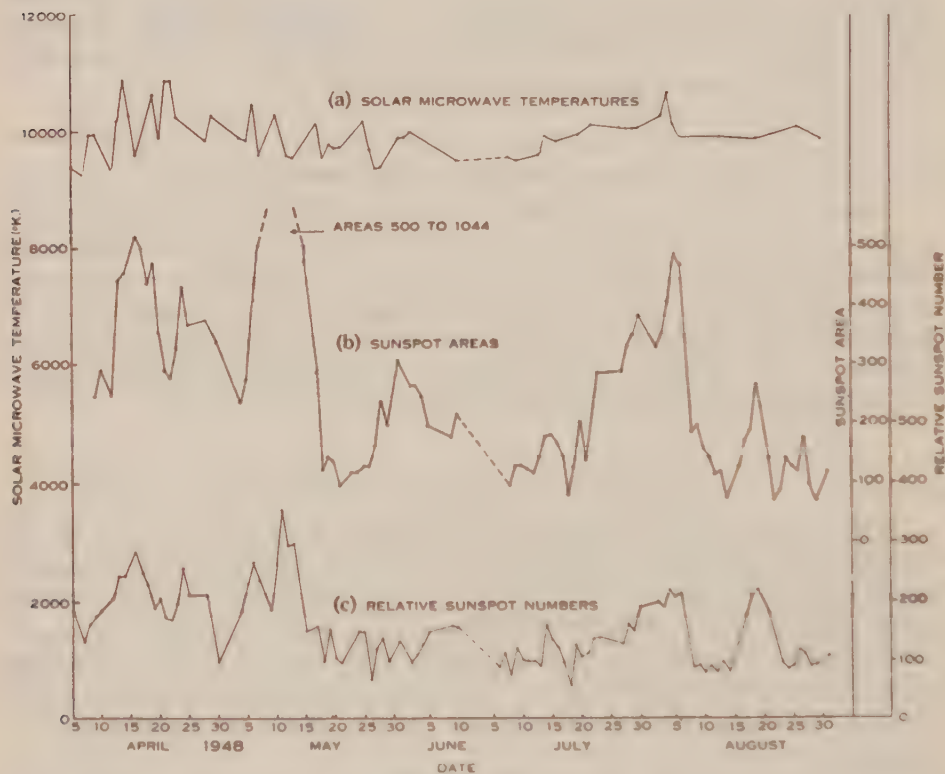


Fig. 1.—Solar microwave temperatures, sunspot areas, and relative sunspot numbers.

(a) *The Normal Level of Radiation*

When corrections were made for atmospheric absorption and variation of the solar diameter from  $32'$ , the amount of radiation received could be converted to an equivalent black-body temperature. The average value of this temperature for all the observations made was  $1.00 \times 10^4$  K., the error being probably less than  $\pm 5$  per cent. The individual results, plotted as daily averages, are shown in curve (a) of Figure 1. This result agrees well with that of Dicke and Beringer quoted above.



The distribution of equivalent temperature about the mean value is shown in Figure 2, where the number of observations which fall into successive small intervals of temperature ( $200^{\circ}\text{K}.$ ) is plotted against the temperature at the centre of the interval. The somewhat irregular form of the histogram is to be anticipated in view of the rather small number (*c.* 50) of observations for this form of analysis.

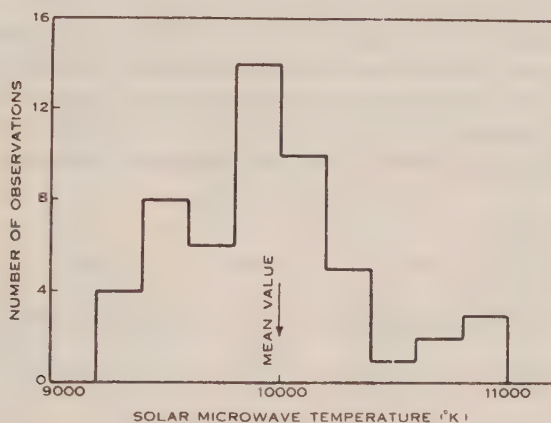


Fig. 2.—Number of daily observations of solar temperature within successive  $200^{\circ}\text{K}.$  intervals.

### (b) Long-Period Variations of Radiation

Most of the daily measurements of solar radiation consisted of two or more readings taken at different times. Thus any short-period fluctuations should tend to be eliminated but the results might be expected to show any variations in the day-to-day general level of radiation. A plot of the daily value of equivalent black-body temperature of the sun is shown in Figure 1. For purposes of comparison a plot is also provided of daily relative sunspot numbers and sunspot areas.\* The period covered is from April 5 to August 30, 1948, with a break between June 11 and July 6.

It will be seen from Figure 1 that the temperature appears to vary from day to day by amounts up to about  $\pm 8$  per cent. These variations are smaller among the later results, being less than  $\pm 3$  per cent. for the period July 14 to August 30 (with one exception, on August 4). The variations were principally due to instrumental errors and decreased as the measuring technique was improved. It is considered that the fluctuation of  $\pm 3$  per cent. obtained during the later period of measurement could be fully accounted for by variations in aerial gain and other factors which could not be checked every day.

\*Kindly provided by the Commonwealth Astronomer.

It is evident that solar radiation at a wavelength of 1.25 cm. remains fairly constant for long periods. There is little or no sign of correlation with sunspot areas or relative sunspot numbers. The single exception is the measurement on August 4 when a rise of about 4 per cent. above the general level was observed. The results at this period showed good consistency and the effect is probably real. On this date large increases of solar radiation were observed at a wavelength of 1.5 m.\* On the other hand, there appears to be no increase in radiation during the period May 6-15 when spot areas rose to more than  $1000 \times 10^{-6}$  of the solar disk.

It is concluded that for the period of the observations solar radiation at 1.25 cm. probably did not vary by more than about  $\pm 3$  per cent. even during periods of solar activity. With a few exceptions, described below, any variations of solar radiation which might exist showed little or no correlation with solar activity as indicated by sunspot areas and numbers or by radio-frequency radiation of longer wavelengths.

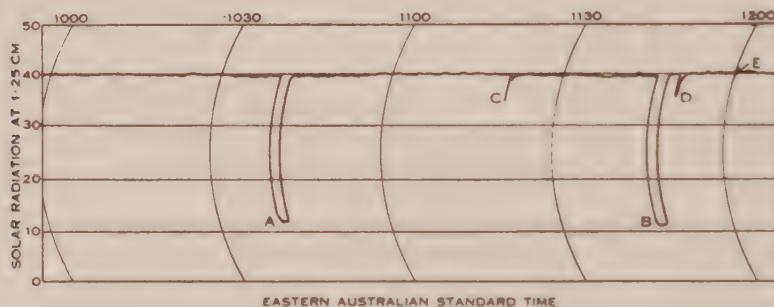


Fig. 3.—A typical record of solar radiation at 1.25 cm. (1000 to 1200 hours Eastern Australian Standard Time, October 5, 1948).

### (c) Short-period Variations of Radiation

On a number of occasions between July and October 1948, continuous records of solar radiation were made at a wavelength of 1.25 cm. Good records could not be obtained if conditions of cloud or wind prevailed, so that the number of such records is limited. In all, recordings were made on 26 days for periods ranging from 2 to 7½ hours.

The best of those records, using the fan-shaped beam described above, showed only slight fluctuations in the received radiation. An example, taken on a day when solar disturbances and ionospheric fade-outs were occurring, is shown in Figure 3. The record plots solar radiation vertically against time from 1000 to 1200 hours Eastern Australian Standard Time. On two occasions (at *A*, *B*) the beam was turned  $5^\circ$  away from the sun so that no solar radiation was received. It will be

\*Records kindly lent by Dr. C. W. Allen of the Commonwealth Observatory.

seen that the total variation of intensity was  $\pm 2$  per cent., which is within the limits of drift due to change of receiver sensitivity, movement of aerial beam, and atmospheric absorption. Deflections at *C* and *D* were due to local electrical interference.

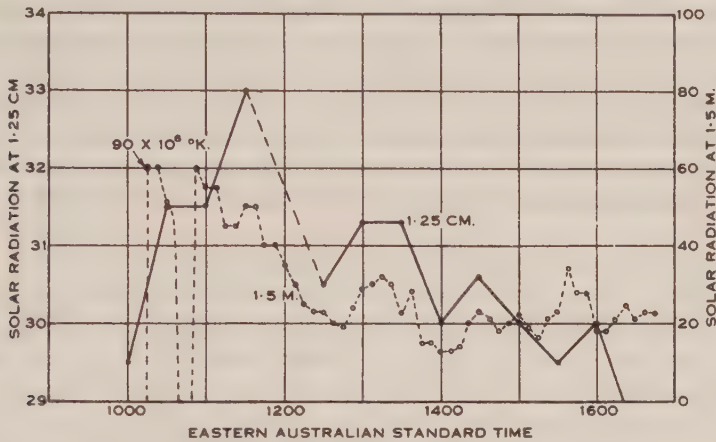


Fig. 4A.—Solar radiation at 1.25 cm. and 1.5 m. averaged over half-hour periods.

On occasions, the curves of recorded intensity showed fluctuations of several per cent., but these could usually be attributed to factors such as movement of the narrow aerial beam. In any case the fluctuations were always so small that their solar origin could be determined only by establishing a correlation between the fluctuations and the other solar phenomena. It has been indicated above that at longer wavelengths fluctuations of solar radiation often correlate with chromospheric flares and with terrestrial ionospheric disturbances. We have, therefore, compared our records with (i) records of chromospheric flares observed at the Commonwealth Observatory, Canberra, and ionospheric fade-outs reported by the Ionospheric Prediction Service of the Department of the Interior, and (ii) 1.5 m. solar intensity records obtained at the Commonwealth Observatory.

Unfortunately, during the period of our observations only one flare and six ionospheric fade-outs were reported. At the onset of one fade-out our records showed an increase of intensity fluctuation of about  $\pm 1$  per cent. full intensity. No additional variation is noticeable in the records during the other fade-outs or the flare. One fade-out occurred at the point *E* on Figure 3 where, if anything, the record is smoother than usual.

A comparison of our records and those for 1.5 m. radiation occasionally shows signs of correlation as in the two examples described below.

On August 6 considerable fluctuations occurred in 1.5 m. solar radiation and fluctuations of about  $\pm 5$  per cent. in the 1.25 cm. record. Portions of the two records are plotted in Figure 4. Long-period fluctuations, averaged for half-hour



intervals, are shown in Figure 4A and short-period variations, averaged for 2-minute intervals in Figure 4B. In each case the scales of intensity are arbitrary.

Another occasion on which there appeared to be a connexion between 1.5 m. and 1.25 cm. radiation was July 29. Unfortunately the 1.25 cm. record for this day was marred by local interference. The 1.5 m. record showed a series of bursts commencing at 1112 hours Eastern Australian Standard Time and finishing at 1128 hours. A flare was observed at 1130 hours. The 1.25 cm. record showed fluctuations commencing at 1107 hours and finishing at 1123 hours, the maximum amplitude of the fluctuations being about  $\pm 5$  per cent. There was no variation in the 1.25 cm. record at the time when the flare was observed. It may be noted that the 1.25 cm. fluctuations commenced and finished about five minutes before the 1.5 m. radiation.

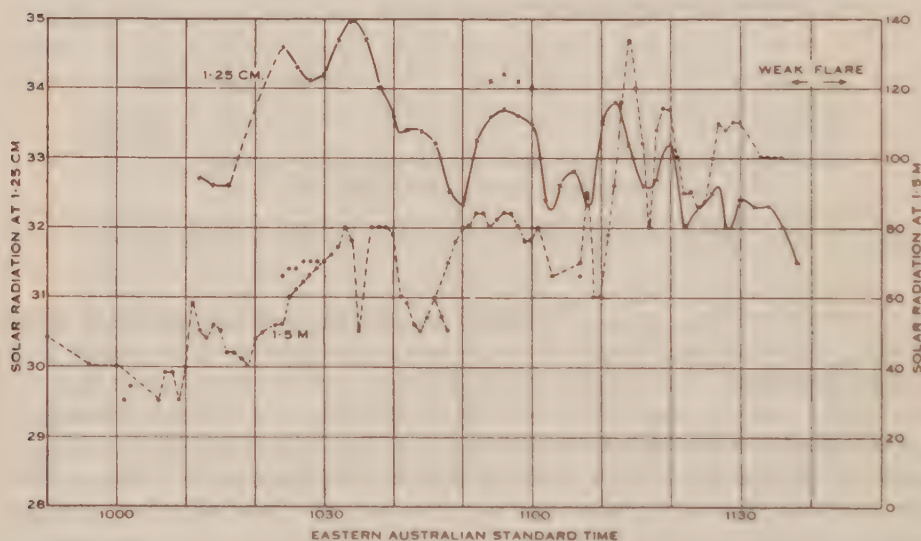


Fig. 4B.—Solar radiation at 1.25 cm. and 1.5 m. averaged over two-minute periods.

The two records described above are the most suggestive of correlation, others for the most part showing little or no such connection. It is concluded that on rare occasions short-period fluctuations of the order of a few per cent. (perhaps  $\pm 5$  per cent.) do occur and that they correlate with large fluctuations in intensity of radiation on a wavelength of 1.5 m.

#### (d) *Distribution of Radiation over the Disk*

When a radiating point source is moved past the aerial diagram provided by the pair of aerials described above, the interference oscillation or fading pattern consists of a series of maxima and zero minima. If the source of radiation is of appreciable size compared with the angular separation of these maxima and minima then the oscillation amplitude is reduced. It is possible, as will be indicated below, to calculate the phase and amplitude of the oscillation corresponding to any radially symmetrical distribution of radiation (as might be expected from the sun).



When the sun moved past the aerial diagram the interference oscillation obtained was found to be a sine curve of small amplitude and phase identical with that from a point source, at the centre of the disk. Since the aerial beam width of each of the two 18-in. reflectors used alone was only about  $1.8^\circ$  to half power and the successive interference minima were about  $0.5^\circ$  apart, only one or two interference oscillations could be used on each run. The amplitude of these oscillations varied considerably, not only because of random receiver noise but also because, as indicated above, the response of each of the two reflectors alone varied throughout each run. Smaller reflectors would have overcome the latter difficulty but the resulting increase in random noise level would more than cancel the gain.

The results of several sets of observations gave oscillations with peak-to-peak value as follows :

3.6, 3.7, 3.0, 4.5, 3.8, 7.2, 0.6, 7.0, 0.7, 4.5, 5.3, 2.0,  
the units being divisions of movement of the recorder pen. The mean value is 3.83 compared with a mean total deflection of 26.0, so that the oscillation amplitude averages 7.4 per cent. of mean deflection. A similar observation was made using a signal generator in place of the sun and driving the aerial beams across the signal generator aerial. The minima were not quite zero, averaging about 5 per cent. of the maxima, presumably because it was not possible to adjust the gain of the two aerals to exact equality. It may easily be shown that this small reduction in amplitude of oscillation will also be present in the same proportion for a distributed source of radiation. The oscillation amplitude due to the sun should, therefore, be increased to 8.2 per cent.

By using graphical integration or by the use of a photoelectric model (which will be described elsewhere) it is possible to estimate the oscillation amplitude from the above interferometer when a radiating source of any size and shape crosses the aerial pattern. It is shown in Appendix I, however, that the response curve is always a sine curve and that if radiation from the sun's disk is symmetrical about a diameter (as we should expect) then the phase of the response curve is such that either a maximum or minimum response is obtained when an aerial lobe is directed at the centre of the sun. The interference technique used is only capable, therefore, of measuring one quantity, namely, the amplitude (either positive or negative) of the oscillation. Thus although any theoretical distribution of radiation may be tested to find if its oscillation amplitude would conform with the observed value for the sun, there are, in general, an infinite number of theoretical distributions which will give the observed oscillation amplitude.

We will compare our results with one (the simplest) distribution of radiation\* which may be made to fit the observations, namely, a uniform disk plus a ring at the limb. It is found that the oscillation amplitude corresponding to a uniformly radiating source of the same angular dimensions as the sun is 15.9 per cent. and that it is in phase with that of a point source at the centre of the sun. The corresponding amplitude and phase for a thin circular ring with diameter equal to the solar diameter is found to be 30.4 per cent. and  $180^\circ$  out of phase with the oscillation due to a point source. It may easily be seen that a composite sun consisting of a uniform

\*The results have been compared(12) with a theoretical distribution and found to be in reasonable agreement.

disk from which 84 per cent. of the radiation was emitted, plus a thin ring from which 16 per cent. was emitted, would give the observed results.

#### IV. CONCLUSIONS

(i) The average equivalent temperature over the solar disk at a wavelength of  $1.25 \text{ cm.}$  is  $1.00 \times 10^4 \text{ }^\circ\text{K.} \pm 5 \text{ per cent.}$  and does not vary appreciably over long periods.

(ii) Day-to-day fluctuations are less than 3 per cent. as compared with several hundred per cent. at wavelengths even as low as  $10 \text{ cm.}$

(iii) Short-period fluctuations do not seem to exceed  $\pm 5 \text{ per cent.}$  even during periods of great solar disturbance. They are usually absent or at least below the level of observational sensitivity at 1 or 2 per cent.

(iv) The distribution of intensity of radiation over the solar disk is consistent with that from a model radiating 84 per cent. of the total radiation from a uniform disk and 16 per cent. from a ring at the limb.

#### V. ACKNOWLEDGMENTS

The work described in this paper was carried out as part of the research programme of the Division of Radiophysics, C.S.I.R.O. Considerable assistance in developing the equipment and making the observations was given by Mr. J. V. Hindman and Mr. G. Trensky.

#### VI. REFERENCES

- (1) SOUTHWORTH, G.—*J. Franklin Inst.* **239**: 285 (1945).
- (2) REBER, G.—*Astrophys. J.* **100**: 279 (1944).
- (3) APPLETON, E. V.—*Nature* **156**: 534 (1945).
- (4) PAWSEY, J. L., PAYNE-SCOTT, RUBY, and MCCREADY, L. L.—*Ibid.* **157**: 158 (1946).
- (5) PAWSEY, J. L., PAYNE-SCOTT, RUBY, and MCCREADY, L. L.—*Proc. Roy. Soc. A* **190**: 357 (1947).
- (6) RYLE, M., and VONBERG, D. D.—*Nature* **158**: 339 (1946).
- (7) GINSBURG, V. L.—*C.R. Acad. Sci. U.R.S.S.* **52**: 487 (1948).
- (8) MARTYN, D. F.—*Proc. Roy. Soc. A* **193**: 44 (1948).
- (9) COVINGTON, A. E.—*Proc. Inst. Radio Engrs.* **36**: 454 (1946).
- (10) DICKE, R. H., and BERINGER, R.—*Astrophys. J.* **103**: 375 (1946).
- (11) PIDDINGTON, J. H., and MINNETT, H. C.—*Aust. J. Sci. Res. A* **2**: 63 (1949).
- (12) PIDDINGTON, J. H.—*Proc. Roy. Soc. A.* (communicated).
- (13) WOOD, R. W.—“Physical Optics.” p. 149. (Macmillan & Co.: London, 1936.)

#### APPENDIX I

Consider two identical aerials separated by a distance  $D$  and having their beams parallel. The resulting interference pattern will have a maximum in the plane which perpendicularly bisects the line joining them. The power received, from a source,

whose direction makes an angle  $\theta$  with this plane, is related to the power  $p_0$  received by either aerial alone (13) by

$$p = 4 p_0 \cos^2 \left( \frac{\pi D \theta}{\lambda} \right)$$

for small values of  $\theta$ . If radiation is received from two equal and independent sources in directions  $(\theta \pm \alpha)$ , the received powers are additive and

$$\begin{aligned} p &= 4 p_0 \left[ \cos^2 \frac{\pi D}{\lambda} (\theta + \alpha) + \cos^2 \frac{\pi D}{\lambda} (\theta - \alpha) \right] \\ &= 4 p_0 + 4 p_0 \cos \frac{2\pi D}{\lambda} \cdot \alpha \cos \frac{2\pi D}{\lambda} \cdot \theta. \end{aligned}$$

It is evident from this equation that the response to any pair of equal radiating sources symmetrically disposed about the plane defined by the direction  $\theta$  will oscillate sinoidally, as  $\theta$  varies with amplitude  $4 p_0 \cos 2\pi D \alpha / \lambda$ . Any surface which has a symmetrical distribution of radiation about the plane  $\theta$  may be considered to be composed of pairs of small, symmetrically placed radiating elements having values of  $\alpha$  and  $p_0$  determined by the distribution. The resultant amplitude of oscillation will be the sum of the individual oscillations which will be in or out of phase according as the angle  $\alpha \lesseqgtr \lambda/2D$ . It is clear that an infinite number of distributions may be found having the same resultant amplitude of oscillation. Thus the intensity distribution of the source cannot be deduced in detail from an observation of the amplitude alone. However, for the values of  $D$  and  $\lambda$  used in the experiment, a small positive (as was observed) or a negative amplitude of oscillation indicates an excess of radiation from near the solar limb as compared to a uniformly radiating disk.

# A REMARKABLE THUNDERSTORM FLIGHT RECORD

By U. RADOK\*

[Manuscript received July 4, 1949]

## Summary

The barogram of a sailplane flight in a thunderstorm at Benalla, Victoria, shows a descent from an altitude of 12,400 ft. to the surface at velocities exceeding 300 ft./sec. over a large part of that height range. The record and the meteorograph used were examined in detail, without definite evidence of instrumental failure. Flight trials gave 100 ft. sec. as maximum descent velocity for the sailplane, while the aerological conditions suggest an upper limit of 130 ft. sec. for the downdraft itself. In comparison with similar data obtained in Florida thunderstorms the lognormal distribution of these data is used to estimate the probability of various large downdraft velocities.

## I. INTRODUCTION

On the last day of the 1948-49 gliding camp held at Benalla by the Gliding Club of Victoria a Grunau Baby II type sailplane piloted by Mr. K. Chamberlin encountered a heat thunderstorm in which it climbed to an altitude of 13,000 ft. above sea-level. The flight has been described in detail by the pilot elsewhere(1); apart from establishing a new Australian altitude record for sailplanes it was remarkable for the rapid descent of the sailplane in the precipitation region of the thunderstorm, where the barogram indicated velocities in excess of 300 ft. sec. over a height range of 6,000 ft. The sailplane in question will not stand up to more than about 160 ft. sec. in a straightout dive; but the pilot reported that the descent was made in a spiral dive with 4 to 5 seconds per full rotation and 60 to 65 m.p.h. indicated air speed, the altimeter indicating a rate of descent exceeding 200 ft. (= 1 scale division) a second. The surface of the sailplane showed local damage due to the impact of hailstones but no evidence of excessive velocities; so that the largest part of the latter must have been contributed by the air surrounding the sailplane.

In view of the significance of large downdrafts for the safety of aircraft and also for the structure of subtropical thunderstorms it seemed of interest to investigate the relevant observational evidence in detail.

## II. METEOROLOGICAL BACKGROUND

The synoptic situation in south-east Australia on January 8 and 9 showed a quasi stationary anticyclone over the Tasman Sea together with a system of tropical depressions over the eastern part of Australia. The latter moved towards the south-east during January 8 and thereby imparted a maritime trajectory to the north-east current dominating Victoria, which originally had continental characteristics.

\*Department of Meteorology, University of Melbourne.



This development was reflected in a number of aerological soundings at Laverton near Melbourne and at Benalla, 120 miles to the north-east. The increase in moisture and heat content was restricted to a layer extending from the surface to about 6,000 ft. at Laverton, and somewhat higher at Benalla ; at higher levels little change occurred during the period in question.

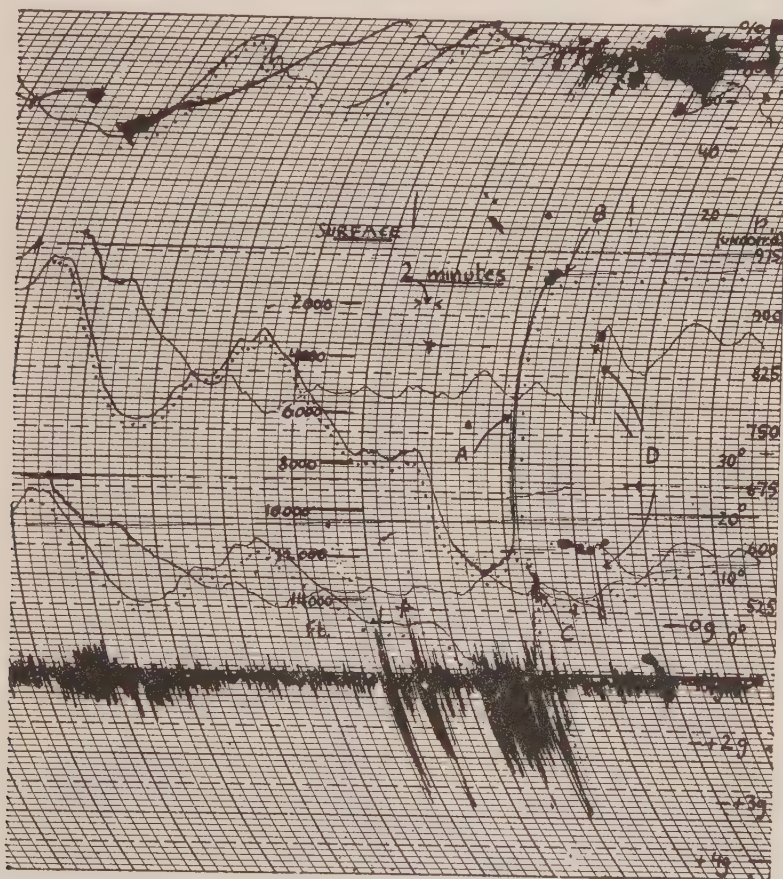


Fig. 1.—Flight record.

The thunderstorm under discussion developed on January 9, 1949, out of a system which had formed around noon over the foothills east and north-east of Benalla. It moved past Benalla Aerodrome at 1530 Eastern Australian Standard Time and was preceded by several large dust devils with very sharply defined trunks visible to about 400 ft. above the surface. Heavy precipitation was visible in its rear but was inadequately recorded by the spare net of rainfall stations ; one station reported almost 1 inch of rain, while others a few miles away had none at all. Three stations recorded falls in temperature of 5-10°C. ; it will be seen below that within the precipitation region the temperature drop was considerably larger.

## III. THE FLIGHT RECORD

The sailplane entered the cloud almost above the aerodrome at 7600 ft.; after climbing to 12,400 ft. above the surface it landed, 25 minutes later, four miles away in the area of precipitation at 1600 hours E.A.S.T. The flight was recorded by a Friez meteorograph attached to the starboard wing strut of the sailplane. The



Fig. 2. — Enlarged view of descent barogram.

chart of this instrument is designed for a record of 2 hours 40 minutes. As the flight in question lasted about four hours its beginning and end appear superposed in the record. This portion of the record is reproduced in Figure 1 with the final

part of the flight underlined by dots. The traces represent (top to bottom) relative humidity, pressure, temperature, and acceleration along the sailplane's vertical axis (positive when directed upwards in straight horizontal flight). Figure 2 gives an enlarged view of part of the pressure trace.

Remarkable features of the record are (i) the very steep descent trace, (ii) its failure to return to the previous surface pressure value (which should have been exceeded, if anything, in the rear of the storm), (iii) the low surface temperature after the landing, (iv) a stoppage of the meteorograph clock earlier in the same flight, and (v) the large positive accelerations accompanying the descent (note that the accelerometer trace has a lag of approximately one minute to that of the barograph).

Points (i) and (ii), together with (v), would seem to suggest a clock failure; however, an isochrone drawn by rotating the barograph pen with arrested clock shows a small but distinct angle with the trace everywhere. Furthermore, there is no kink in the pressure trace, and also the pilot's report quoted in Section I, together with points (ii) and (iii) above, indicate that the descent actually took place very rapidly.

#### IV. THE EVIDENCE OF THE BAROGRAM

Before evaluating the barogram the pressure corrections for various stages of the flight were determined by means of a flight similitude test in a pressure chamber. This test was made at room temperature and showed only deviations of a few millibars at all pressures. The starting pressure was reproduced exactly in this case, which suggested that the lag observed in the flight record represented a temperature effect. When testing the temperature compensation it was found that the instrument had developed a large temperature coefficient, owing to the incorrect replacement of a bimetallic shaft after the fitting of the accelerometer unit.

With the corresponding corrections the height scale in Figure 1 was determined; these heights refer only to the ascent, and are minimum values, since in their calculation it has been assumed that the barometric cell had the temperature of the well-exposed bimetallic strip. Thus the surface pressure recorded after the landing corresponds to a temperature of about 0°C. of the barometric cell, and is an indication of a rapid descent.

Next, to obtain the vertical velocities from the barogram the usual way is to measure both height differences and the corresponding time intervals. This method requires, however, steady conditions for periods of one minute or more; and even then would be inaccurate owing to the small scales of the record. On the other hand, both the pressure and the time scales are such that even small differences in the vertical velocity will give rise to distinct changes in the angle of the trace with the isobars. Thus that angle will be an accurate measure of the vertical velocity, but it will for the same velocity also vary with pressure.

The relation between the vertical velocity and the slope of the barograph trace



is illustrated in Figure 3 and has the simple form

$$1/v' = \cot \delta - \cot \omega,$$

where apart from a scale factor  $v'$  is the time rate of pressure change (mb. 'sec.), while  $\delta$  and  $\omega$  are the angles of inclination of the trace and the isochrones, respectively. The latter can be regarded as circles, and thus their slope is known as function of their radius, the actual pressure,  $p$ , and the pressure of the isobar through their centres,  $p_0$ .

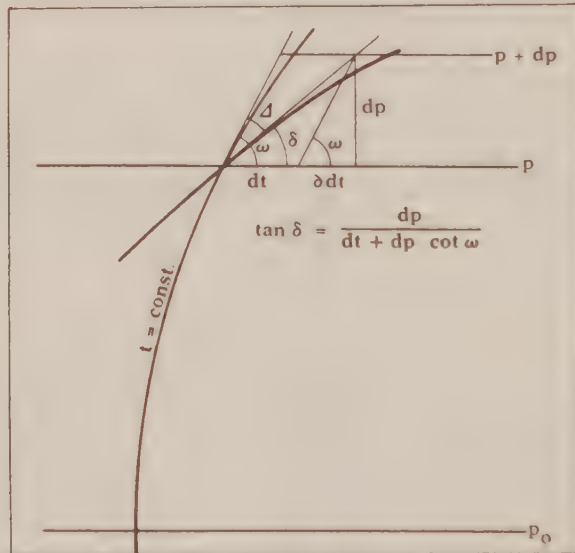


Fig. 3.—Slope of diagram trace and rate of pressure change.

The variable  $v'$  is converted into the actual vertical velocity by means of the "barometric height interval," i.e. the vertical distance corresponding to a change in pressure of 1 mb. This is given by

$$1 \text{ mb. (m)} = (7991/p) (1 + 0.04 t)$$

for the pressure  $p$  in mb. and the temperature  $t$  in °C. at any level.

In the present case a further adjustment is necessary to take into account the temperature coefficient of the instrument used. This has been done subsequently, however, and Figure 4 thus shows the lines of equal vertical velocity, as a function of pressure and the slope of the barogram, for a normally functioning Friez instrument. In this diagram the isotaches crowd together for velocities above 20 m. sec. which renders accuracy impossible; however, even for larger velocities the slope method retains the advantage of not requiring straight sections in the trace.

In Figure 4 the crosses show the evaluation of the descent recorded in Figure 1. After applying the necessary corrections the height-velocity diagram was converted into the graphs for height, velocity, and acceleration on a time scale shown in Figure 5.



The validity of Figure 5 is dependent on the accurate functioning of the meteorograph clock which will be examined in Section VI. It should be noted here that the shape of the diagrams in Figure 5 does not alter greatly with changes in the maximum velocity; i.e. when 80 m./sec. is substituted for 100 m./sec. in Figure 5 the total time of the descent is increased by only 3.5 seconds, while the peak accelerations become  $-1.4$  g and  $+1.7$  g, respectively.

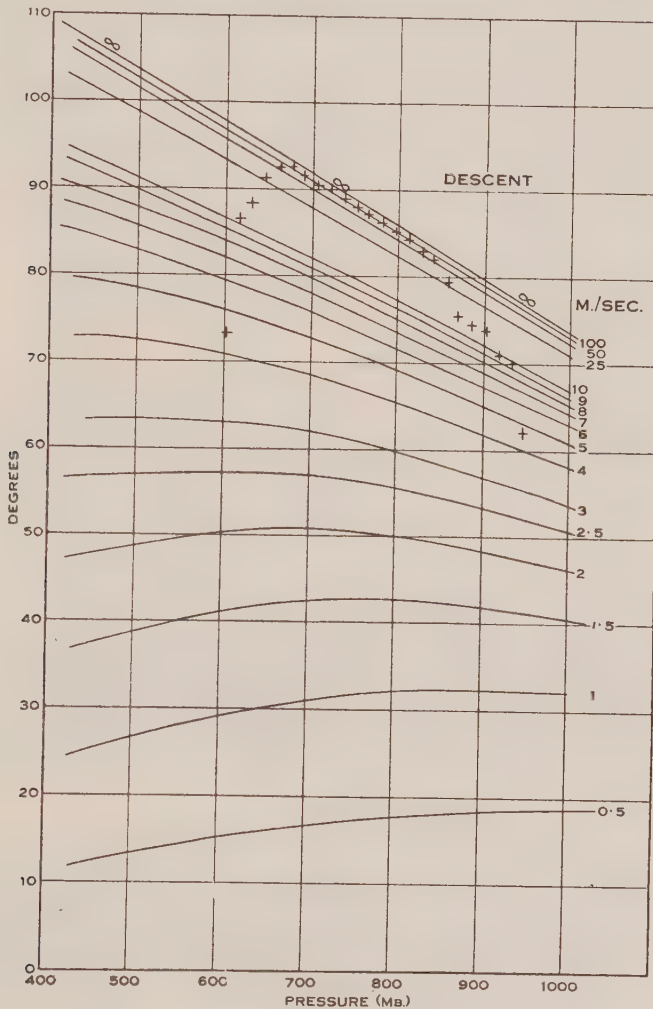


Fig. 4.—Velocity diagram for Friez meteorograph (descents).

## V. THE EVIDENCE OF THE FLIGHT TRIALS

From the velocities found in the preceding section that of the descending sailplane with respect to the surrounding air has to be deducted, in order to arrive at the actual downdraft velocity. The attitude of the sailplane during the descent

was already suggested by the pilot's report as that of a spiral dive; this explained the fact that the recorded accelerations are opposite in sign to those expected from Figure 5. Thus the check on the accuracy of the vertical velocities computed, which otherwise would have been provided by the independent acceleration record, could not be applied.

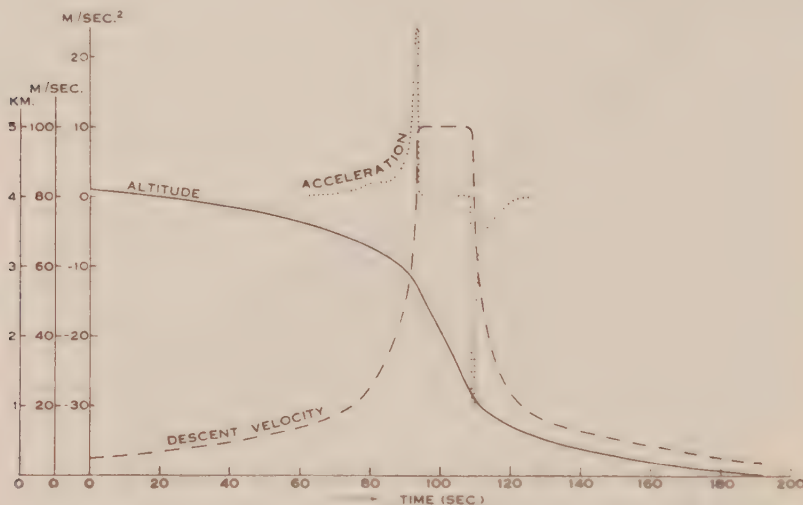


Fig. 5.—Reconstruction of descent.

The probable descent velocity of the sailplane itself was determined from three flights (piloted by Mr. N. Hyde) in which the sailplane was towed to an altitude of 6000 ft. and made to descend by holding the control stick in its rearmost position. The only stable mode of descent under these conditions appeared to be a steep slipping spiral. The descent velocity was evaluated both from barograms and from observations taken with two balloon theodolites, and was found to be 8 m. sec. in two cases, but in excess of 20 m./sec. in the third case.

This discrepancy suggests that the reproduction of the relevant mode of descent was not necessarily achieved in these trials; this is confirmed also by the lower accelerations recorded during the trials (2.3 g maximum). Thus there seems sufficient reason to allow for a velocity of the sailplane itself of up to 30 m. sec. This would leave a net downdraft of the order of 60-70 m. sec. (200-230 ft. sec.) on the evidence of the barogram evaluation.

## VI. THE EVIDENCE OF THE CLOCK

Thus even when extreme values are assumed for the sailplane's own descent velocity the resulting downdraft is far in excess of any previously recorded. Before looking to the aerological conditions for its justification it seemed therefore advisable to give some attention to the meteorograph clock. A complete stoppage of the

latter was excluded by the appearance of the trace. There remained, however, the possibility that the motion of the clock had been affected by the high accelerations during the descent.

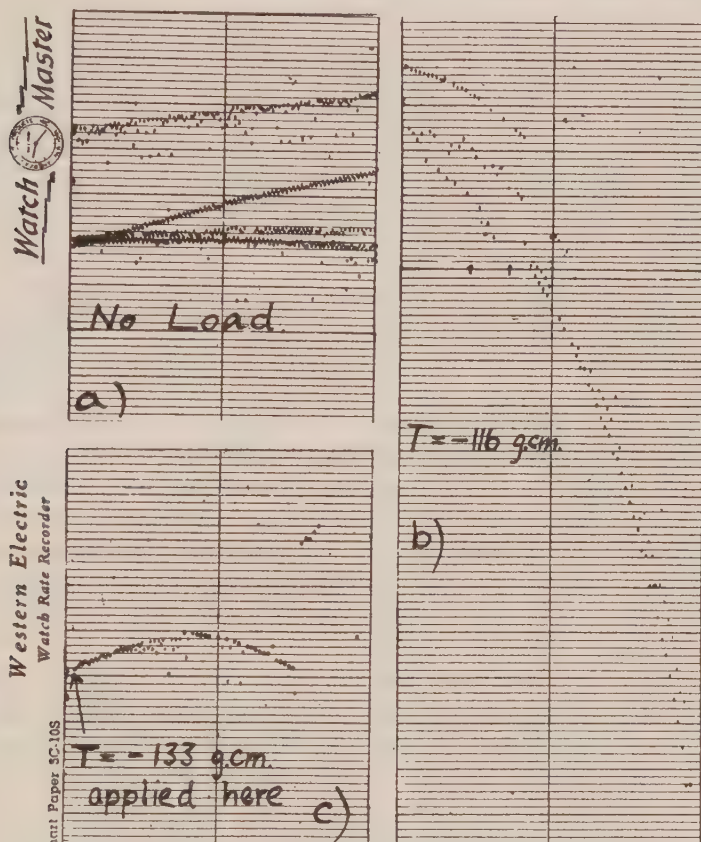


Fig. 6.—Results of clock test.

This suspicion gained support, when it was found that the steel clip holding the chart on the recorder drum showed an unbalanced excess of weight. Thus, with the drum in the position it occupied during the descent, a torque was exerted opposing that of the clock spring. This torque would increase from 18 g. cm. at normal gravity to 63 g. cm. at the maximum acceleration during the descent; in fact, since only a component of those accelerations was recorded the countertorque may have been even higher.

The motion of the clock under different loads was tested by means of a watch rate recorder. This consists of a drum revolving five times a second, on which the sounds of the escapement are electronically recorded. Since the correct speed calls for five movements of the lever per second it will be recorded as a straight horizontal

line across the chart. Positive slope of the record indicates fast motion, negative, slow. The width of the chart is covered in 30 seconds, and one vertical division represents a gain or loss of five seconds in 24 hours.

Some of the records obtained for the meteorograph clock are shown in Figure 6. Without external load the speed was practically correct and constant for various attitudes of the clock (Fig. 6a). Small counter-torques, applied statically, had no visible effect. For larger loads a small decrease in speed was found, which on further increase of the load developed into an increasing retardation (Fig. 2b). Finally an even higher load arrested the motion of the clock altogether within less than 30 seconds after being applied (Fig. 6c; note the series of fast beats immediately prior to stopping; these were probably due to a balance wheel oscillation of reduced amplitude and in one case were actually obtained as a quasi-steady condition, for a load between those of Figs. 6b and 6c).

Of particular interest here is the parabola-shaped record of Figure 6b which indicates a progressive decrease, at a constant rate, in the speed of the clock. The rate of the decrease was computed from the time required by the slope of the parabola to change from zero (= correct speed) to that corresponding to a loss of 24 hours during the test (= no motion). This time was found to be about 32 minutes.

The conclusion is that, even if such an extrapolation were permissible for as much as three minutes, the loss in speed would have remained below 10 per cent. during that period and thus would not have affected the slope of the descent trace to any great extent. In addition it must be noted that the test values of the counter-torque required to produce the effects of Figures 6b and 6c were much larger than those estimated from the unbalanced weight on the drum circumference and the accelerations during the descent.

Nevertheless it cannot be ruled out completely that in effect the motion of the clock was slowed down through a number of short stoppages during at least part of the descent. That may be the explanation for the wavy appearance of some sections of the trace. The effect on the slope of the barogram could of course be serious.

## VII. DOWNDRAFT VELOCITIES SUGGESTED BY THE AEROLOGICAL CONDITIONS

The upper air temperatures and relative humidities observed on January 9 above Benalla are shown in Figure 7 plotted on a Tephigram. The morning flight was recorded by a different meteorograph and at a flying speed higher than that of the sailplane; the highest point of this flight coincided with the top of a marked haze layer indicating the upper boundary of the moist air. The difference between the morning and afternoon temperatures must be attributed partly to the defect in the sailplane meteorograph discussed in Section IV. The same consideration precluded a strict comparison of the afternoon temperatures with those given by the radiosonde flight at 1830 hours E.A.S.T. on January 9 at Laverton, which for that reason have not been included in Figure 7; above the 800 mb. level the Laverton temperatures were higher than those recorded earlier by the sailplane at Benalla (which is consistent with the presence of Easterlies), but the average difference of 4°C. appears too large. As already mentioned in Section II the radiosonde flight indicated rather dry air above 800 mb., with moist adiabatic lapse rate above 650 mb.



To estimate possible downdraft velocities temperature observations would be required both for the downdraft itself and for corresponding levels in the air outside the cloud. In absence of these it seems still possible to arrive at some rough estimate by means of plausible assumptions regarding the nature of the downdrafts and the conditions in the dry upper air.

The environment of convective clouds is of fundamental importance for their development, according to recent work (2) which showed that part of it tends to be entrained into such clouds. By becoming saturated the air of the environment can, at the most, be cooled down to its wet-bulb temperature. As a result of its mixing with the cloud air the lapse rate in the cloud becomes greater than the moist adiabatic when the cloud air is ascending through a colder environment.

This effect is clearly indicated in the sailplane ascent curve in Figure 7 by the steep lapse rate and the low values of the relative humidity in the cloud. This allowed extrapolation to the temperatures in the uppermost part of the flight (shown by the broken line in Figure 7), where the acceleration record obliterated that of the temperatures (cf. Fig.1).

The rate of temperature decrease with altitude for the cloud as a whole cannot have been less than that given by the moist adiabat marked by dots in Figure 7, approximately in continuation of the sailplane ascent sounding. For the environment above the moist layer the radiosonde flight at Laverton suggests the same lapse rate; while the temperatures in the environment could have been only little higher than those at corresponding altitudes in the cloud (the difference in density between the dry outside air and the moist cloud air leading, in that case, to the rise of the latter); nor could they have been much lower, in view of the already large temperature difference between Benalla and Laverton.

As an approximation therefore the right-hand dotted curve in Figure 7 can be taken to represent both the stratification of the (saturated) cloud air and that of the environment (with negligible humidities). There exists no evidence in the present case justifying the assumption that the cloud was much warmer than its surroundings, as has been frequently found elsewhere (temperature differences of up to  $15^{\circ}\text{C}.$ , cf. 3).

This leaves the nature of the downdrafts to be determined. Now since the mixing of the cloud air with that of the environment is often far from being complete (4) large parcels of the latter could conceivably be cooled to their wet-bulb temperature by evaporation from rain and thus become much heavier than the surrounding cloud air. If the entire downdraft consisted of such air its temperature at different heights would be given, at any rate in the dry layer, by the left-hand dotted curve in Figure 7, which represents the wet-bulb temperatures of the hypothetical environment above 720 mb. In the moist layer below that level entrainment would no longer have a cooling effect, rather the opposite; this has been neglected here, however, so that the hypothetical temperature curve for the descending air is made to follow a moist adiabat to the surface.

The conditions assumed for the purpose of estimating the highest possible downdraft velocities are thus extremely biased towards the largest temperature differences (and velocities) compatible with the incomplete observational evidence.

Under such circumstances it must be described as remarkable that the surface temperature recorded after the landing (marked by a cross in Figure 7) lies close to that expected from the (independently derived) hypothetical descent.

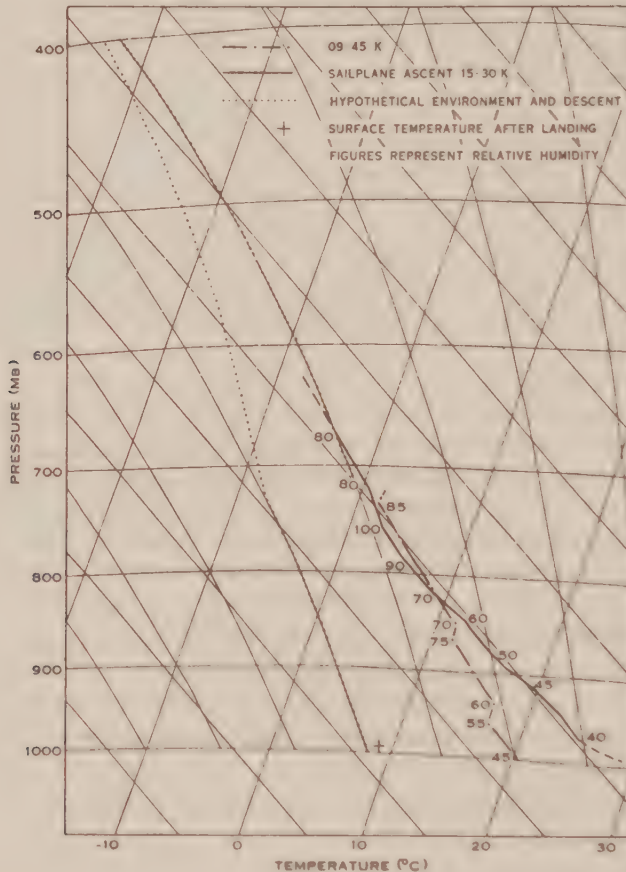


Fig. 7.—Aerological conditions.

The velocities reached during that descent can be computed from the two dotted curves by means of Bernoulli's equation. The energy losses involved in such a descent have been estimated by Raethjen(5) to amount to about 50 per cent. In the present context a rough upper limit to the possible velocities should be obtained by neglecting all losses. Higher velocities could of course arise in compensation of ascending motion elsewhere, but this cannot be substantiated in the present case.

For the descent merely under the influence of the Archimedean forces Bernoulli's

equation then takes the simple form

$$v^2 = v_0^2 + 2g \int_z^{z_0} \frac{T - T'}{T} dz,$$

where  $v$  is the velocity at the level  $z$ ,  $v_0$  that at the starting level  $z_0$ , and  $T$  and  $T'$  are, respectively, the absolute temperatures (strictly speaking : virtual temperatures) inside and outside the cloud.

The following table gives the velocities at various levels for a parcel of air that started its descent at the 400 mb. level with negligible velocity  $v_0$ . Such a descent could be initiated in compensation of a rising current, or after the penetration of such a current into warmer surroundings.

TABLE 1  
AEROLOGICAL ESTIMATE OF DOWNDRAFT VELOCITIES

Altitude (km.)	7	6	5	4	3	2
Velocity (m./sec.)	7	16	24	31	39	47

In the computation of the above velocities not only all losses have been left out of account, but also certain accelerating factors such as falling rain or hail. Thus the upper limit for the velocity of descent could reasonably be drawn at 40 m./sec. at the 3 km. level.

#### VIII. CONCLUSIONS AND COMPARISON WITH FLORIDA THUNDERSTORM DATA

The combined evidence of Sections IV-VII thus cannot be completely reconciled. The quality of the record (Section IV) and various suspicious circumstances (Section VI) render it inadvisable to base a claim for velocities in excess of all previously encountered on this one observation. There remain, however, definite indications that an unusually strong downdraft existed in the precipitation region of the thunderstorm in question.

The likelihood of the velocity claimed can finally be judged, to some extent, by bringing it into comparison with similar observations made in the course of the United States Thunderstorm Project at Florida(6). Table 5 (p. 13) of reference (6) gives the frequencies of different rates of vertical displacement observed on horizontally flying aircraft in thunderstorm clouds. Such frequencies are as a rule plotted in cumulative form on a logarithmic scale, against the vertical velocity on a linear scale. In the present case the logarithmic frequency scale was replaced by a normal

probability graduation; this has the effect that data following a normal distribution will be plotted on a straight line.

The above leaves open the choice of a suitable velocity scale. For the data of reference (6) it was found that a straight line plot resulted from choosing a logarithmic scale for the abscissa (vertical velocity).

Figure 8 thus shows the interesting fact that the logarithms of the vertical velocities observed in a number of Florida thunderstorms can be considered as normally distributed. Extrapolation to the order of magnitude suggested by the Benalla flight record indicates that on the strength of the Florida data, downdrafts of the order of 30-35 m./sec. (100-110 ft./sec.) seem not altogether unlikely (1 chance in 10,000); however, velocities of 55 m./sec. (180 ft./sec.) and more must be classed as very improbable (1 chance in 1,000,000 and less), if conditions in the south-east of the United States and south-east Australia are at all comparable.

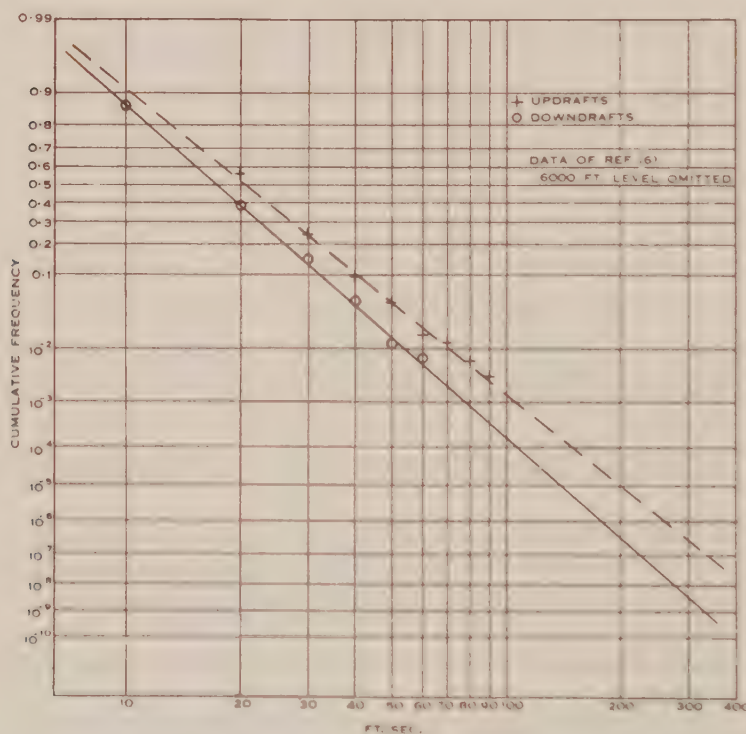


Fig. 8.—Distribution of Florida data.

## IX. ACKNOWLEDGMENTS

The above investigation arose in the course of a programme of research work on gusts, which was undertaken, with the support of the Aeronautical Research Laboratories, Department of Supply and Development, and of the Gliding Club of Victoria, on behalf of the Flight and Ground Loads Panel of the Conference for Operational



Research. Many persons were consulted at various stages, so that it is impossible to acknowledge in detail all the help received. The author desires, however, to thank all members of the Gliding Club of Victoria, the Aeronautical Research Laboratories, the Section of Meteorological Physics, C.S.I.R.O., the Commonwealth Meteorological Bureau, the Australian Army Signal Corps, and the Royal Australian Air Force, who assisted him in many difficulties. Helpful advice was also received from Dr. F. Loewe and Professor W. Georgii.

#### X. REFERENCES

- (1) CHAMBERLIN, K.—*Sailplane and Glider* 17: 27 (1949).
- (2) STOMMEL, H.—*J. Met.* 4: 91 (1947).
- (3) HOEHNDORF, F.—F.I.A.T. Rep. No. 1064, p. 55 (1947).
- (4) BYERS, H. R., and BRAHAM, R. R.—*J. Met.* 5: 71 (1948).
- (5) RAETHJEN, P —“Einführung in der Physik der Atmosphäre.” Vol. 2 (B. G. Teubner: Leipzig, 1942.)
- (6) BYERS, H. R., ET AL.—U.S. Weath. Bur. Techn. Pap. No. 7, April 1948.

# BOND LOCALIZATION ENERGY

## I. DEFINITION, METHODS OF COMPUTATION, AND RELATION TO BOND ORDER

By R. D. BROWN\*

[Manuscript received June 6, 1949]

### Summary

A new quantity—the bond localization energy—which has some advantages in the theoretical discussion of the chemical reactivities of various bonds, is introduced. Its relationship to the activated complex is indicated. It is shown to be expressible in terms of the resonance energies of certain molecules related to the molecule being considered. The various methods available for computing resonance energies by the LCAO method are outlined and some of these are used to derive numerical values for the localization energies of particular bonds.

The relationship between bond localization energy and mobile order is investigated, and the significance of the relations thus found is indicated.

## I. INTRODUCTION

From time to time in the LCAO molecular orbital treatment of organic molecules a number of auxiliary concepts has been introduced. Thus the quantities electron densities(1), mobile bond orders(2), resonance energies(3, 4), free valences(5), atomic polarization energies(6), mutual atom and bond polarizabilities(7), and conjugating powers(8) have been discussed and their relationships to various chemical and physical properties of the molecule have been elaborated. The purpose of the present paper is to introduce a new quantity—the *bond localization energy*—which appears to have a number of advantages over comparable quantities already in use. The method of computing the bond localization energy will be outlined and its relationships to other quantities will be discussed. The application of the bond localization energy to the interpretation of various chemical reactions will be dealt with in a subsequent paper.

## II. DEFINITION

The localization energy of a given bond in a molecule is defined as the energy required to perturb the molecule so that the bond in question becomes a double bond, i.e. so that two of the  $\pi$ -electrons are localized in the bond under consideration. A similar type of perturbation of the molecule has been suggested by Kooyman and Ketelaar(9) in a discussion of the ozonolysis of naphthalene. They calculated the localization energies for some of the bonds of naphthalene using experimental resonance energy data. Their procedure suffers from two disadvantages. Firstly it relies upon the existence of the relevant resonance energy data, and secondly it is incorrect to assume, for example, that the resonance energy of the system *o*-benzo-

\*Chemistry Department, University of Melbourne.

quinodimethane is the same as that of octatetraene; in fact, calculation shows that the former has the greater resonance energy by about 6 kcal. mol.<sup>-1</sup>. Both of these defects may be avoided by basing the energy computations on the LCAO method. From the recent discussion by Dewar(10) it seems that by means of this latter method we may derive resonance energies which are at least as reliable as the best available experimental data.

The bond localization energy may be computed in a manner similar to that used by Wheland(6) for the calculation of atomic polarization energies. In analogy to his nucleophilic and radical atomic polarization energies, the energies required to perturb the molecule so that one particular bond becomes a single bond or a bond containing one  $\pi$ -electron could easily be calculated. However, it does not seem likely that such perturbations would be involved in any of the common chemical reactions of aromatic bonds.

### III. METHOD OF COMPUTATION

As is usual in theoretical considerations of extended conjugated systems, we shall consider only the  $\pi$ -electrons. The phrase "a molecule containing  $n$  carbon atoms" will be used synonymously with the more precise phrase "a molecule containing a conjugated system of  $n$  carbon atoms"; it does not seem likely that the interactions of  $\sigma$ - and  $\pi$ -electrons will affect the results of the present theoretical considerations appreciably. The total  $\pi$ -electron energy in the unperturbed molecule will be represented by  $\mathcal{E}$  and the energy of the  $\pi$ -electrons in the perturbed molecule by  $\mathcal{E}^\dagger$ . Such a perturbed molecule would be expected to represent approximately

the state of the molecule in the activated complex of a reaction such as ozonolysis. If it is the bond joining atoms  $a$  and  $b$  which is so localized, then the  $\pi$ -electron energy of the activated complex may in good approximation be represented as

$$\mathcal{E}_{ab}^\dagger = \mathcal{D} + \mathcal{E}_{ab}, \quad \dots\dots\dots(1)$$

where  $\mathcal{D}$  is the energy of the  $\pi$ -electrons in ethylene and  $\mathcal{E}_{ab}$  is the  $\pi$ -electron energy of the residual molecule derived from the original molecule by removing atoms  $a$  and  $b$  from the conjugated system.

If the original molecule contains a conjugated system of  $2n$  carbon atoms (and thus  $2n$   $\pi$ -electrons) then, from the definition of the resonance energies,  $R$ ,  $R_{ab}$ , of the two conjugated systems under consideration, we have

$$\mathcal{E} = n\mathcal{D} - R, \quad \dots\dots\dots(2)$$

$$\mathcal{E}_{ab} = (n - 1)\mathcal{D} - R_{ab}, \quad \dots\dots\dots(3)$$

and it will be noticed that  $R$  and  $R_{ab}$  are both positive quantities. In some cases the removal of atoms  $a$  and  $b$  from the molecule will leave two residual molecules. This causes no difficulty because then  $R_{ab}$  is just the sum of the resonance energies of the two residues. However, in some cases the two residual molecules each contain an odd number of carbon atoms, i.e. they are radicals. In the latter event equation (3) is still valid if we define the resonance energy of a free radical in the appropriate

manner. Thus a radical composed of  $2n + 1$  carbon atoms is assumed to contain  $2n + 1$   $\pi$ -electrons, and the resonance energy is defined as  $(2n + 1)$  times the energy of a  $\pi$ -electron in ethylene minus the total  $\pi$ -electron energy of the radical. One further complication arises when one or both of the radicals is a non-alternant(11), e.g. *cyclopentadienyl*. This point will be considered later.

Now the bond localization energy,  $B_{ab}$ , of the bond joining atoms  $a$  and  $b$  is defined by

$$B_{ab} = \mathcal{E}_{ab}^\dagger - \mathcal{E}, \quad \dots\dots\dots(4)$$

and from equations (1), (2), and (3), this may be written

$$B_{ab} = R - R_{ab} \quad \dots\dots\dots(5)$$

Consequently the problem of the computation of the bond localization energy reduces to the task of computing resonance energies for the original molecule and the residual molecule or molecules.

#### IV. METHODS OF CALCULATING RESONANCE ENERGIES

It is convenient to employ the LCAO molecular orbital approximation for the resonance energy calculations (but the valence bond method would serve equally well). In the LCAO method the secular equations for a given molecule are well known to be\*

$$\sum_{j=1}^{2n} c_j (H_{jk} - ES_{jk}) = 0, \quad k = 1, 2, \dots, 2n, \quad \dots\dots\dots(6)$$

where the conjugated system is comprised of  $2n$  carbon atoms (for the present we shall restrict our attention to hydrocarbons). Proceeding then upon the following assumptions :

- (i) that all of the coulomb integrals,  $H_{jj}$ , are equal ;
- (ii) that the overlap integrals,  $S_{jk}$ ,  $j \neq k$ , are all zero ;
- (iii) that the exchange integrals,  $H_{jk}$ ,  $j \neq k$ , are zero unless  $j$  and  $k$  are adjacent atoms ;

the equations (6) are reduced to

$$c_k(\alpha - E) + \sum_j c_j \gamma_{jk} = 0, \quad k = 1, 2, \dots, 2n, \quad \dots\dots\dots(7)$$

where the summation is over all atoms,  $j$ , which are adjacent to  $k$ ,  $\alpha$  has been written for  $H_{jj}$ ,  $\gamma_{jk}$  for  $H_{jk}$ , and  $E$  is the energy of the system. On the further assumption that all of the  $\gamma_{jk}$  are equal†, division through each equation by  $\gamma$  gives the result

$$c_k x + \sum_j c_j = 0, \quad k = 1, 2, \dots, 2n \quad \dots\dots\dots(8)$$

in which the substitution

$$x = (\alpha - E)/\gamma \quad \dots\dots\dots(9)$$

has been made. The equations (8) may be reduced to a polynomial equation in

\*For a general account of the method see (7).

†The energies may be computed with little further labour without making this assumption (12), but except for molecules containing bonds with very widely varying mobile order,  $p$ , the difference is very slight.



$x$  of degree  $2n$ , which can be shown to give  $2n$  real roots; we shall represent these roots by  $x_r$ ,  $r = 1, 2, \dots, 2n$ . To each of these roots there corresponds a  $\pi$ -type molecular orbital of energy  $E_r$  such that (cf. equation (9))

$$E_r = \alpha - x_r \gamma, \quad \dots\dots\dots(10)$$

and because both  $\alpha$  and  $\gamma$  are negative, the molecular orbitals of lowest energy will be those corresponding to the most negative  $x_r$ . It is thus convenient to number the roots so that  $x_1$  will correspond to the lowest energy and the others are numbered in order up to  $2n$  which corresponds to the highest energy. In the ground state of the molecule the  $n$  lowest energy levels are occupied, each by two electrons, and the other levels are unoccupied. Thus it is easy to see that

$$\mathcal{E} = 2n\alpha - 2\gamma \sum_{r=1}^n x_r \quad \dots\dots\dots(11)$$

$$\mathcal{D} = 2\alpha + 2\gamma \quad \dots\dots\dots(12)$$

so that from (2), (11), and (12), we arrive at the result

$$R = 2(n + \sum x_r) \gamma \quad \dots\dots\dots(13)$$

Regarding now the bond localization energy,  $B_{ab}$ , as a quantity arising in the LCAO molecular orbital theory, it is convenient to define it, in view of (5) and (13), as

$$B_{ab} = 2(1 + \sum x_r - \sum x_r^\dagger) \gamma \quad \dots\dots\dots(14)$$

It is now necessary to consider the validity of equation (14) when the removal of atoms  $a$  and  $b$  results in two residual molecules. It seems reasonable to assume that when the molecule becomes perturbed, the remaining  $(2n-2)\pi$ -electrons will occupy the  $n-1$  orbitals of lowest energy; further it is probable that this will also be the case when two residual molecules or free radicals are formed, even if this results in an uneven distribution of  $\pi$ -electrons between the two\* (i.e. a distribution other than just one  $\pi$ -electron per carbon atom). The work of Coulson and Rushbrooke(11) shows that, when the residues are both alternants, the sharing of  $\pi$ -electrons will always be even because for such molecules the unoccupied orbitals all have positive  $x_r$ , and the singly occupied orbital for alternant free radicals always corresponds to  $x_r = 0$ . However, when one or more of the residual molecules or free radicals is a non-alternant it can happen that† some unoccupied orbitals correspond to *negative*  $x_r$ . If such an unoccupied orbital is of appreciably lower energy than an occupied orbital in the other residual molecule which is normally occupied, then a more stable system will result if the last pair of  $\pi$ -electrons occupy the former orbital (thus leaving the latter residual molecule with less than its quota of  $\pi$ -electrons). Thus the definition (14) is preferable to (5) provided the summation over the  $x_r$  applies to those  $n-1$  orbitals of lowest energy regardless of their distribution

\*The LCAO approximation does not take into account the mutual repulsions of  $\pi$ -electrons and so makes no allowance for the energy required to produce this charge separation. Consequently this separation of charge will only occur when the resultant gain in energy due to occupancy of the new orbital outweighs this.

†Fulvalene for example has an unoccupied orbital with  $x = -0.25410$ .

over the residual molecules. It is hardly likely that the difference from the result derived from resonance energy data would be large enough to be important in qualitative discussions\*. Such a difference will be illustrated in a subsequent paper in the case of vinylfulvene.

Returning for the moment to assumption (ii) which was used in deriving equations (7) from equations (6), it may be shown by direct computation of the integral  $S_{jk}$  using normalized Slater-Zener atomic orbitals with  $Z' = 3.23$  that its magnitude when  $j$  and  $k$  are separated by  $1.39 \text{ \AA}$  is  $0.25$ . The possibility of computing the  $\pi$ -electron energy without assuming, as had formerly always been done, that  $S = 0$  was indicated by Mulliken and Rieke(13), and the calculation of resonance energies using  $S = 0.25$  has been carried out by Wheland(4). If we now replace assumption (ii) by the assumption that  $S_{jk}$  is finite when  $j$  and  $k$  are neighbours, otherwise zero, then equations (6) become

$$c_k(\alpha - E) + \sum_j c_j(\gamma - SE) = 0. \quad k = 1, \dots, 2n, \quad \dots(15)$$

where all the  $\gamma$  have been assumed equal, and  $S$  has been written for the overlap integral between neighbouring atoms. Again equations (15) reduce to equations (8) upon dividing through by  $\gamma - SE$  and making the substitution

$$x = \frac{\alpha - E}{\gamma - SE} \quad \dots(16)$$

Equation (16) is easily rearranged to give

$$E'_r = \alpha - \frac{x_r}{1 - x_r S} \beta, \quad \dots(17)$$

which is the expression for the energy,  $E'_r$ , of the  $r$ th molecular orbital in terms of the root  $x_r$  of the secular equation, and  $\beta$  has been written for  $\gamma - Sx$ . This may conveniently be written

$$E'_r = \alpha - e_r \beta \quad \dots(18)$$

in analogy to equation (10), and similarly in analogy to equations (11), (12), and (13) we find (putting  $S = 0.25$ )

$$\mathcal{E}' = 2n\alpha - 2\beta \sum_{r=1}^n e_r, \quad \dots(19)$$

$$\mathcal{D}' = 2x + 1.6\beta, \quad \dots(20)$$

$$R' = 2(0.8n + \sum e_r)\beta, \quad \dots(21)$$

and again we arrive at the definition of bond localization energy as

$$B'_{ab} = \{1.6 + 2(\sum e_r - \sum e_r^\dagger)\} \beta \quad \dots(22)$$

It will be noticed that the corresponding quantities for the case of non-zero overlap integral have been distinguished by dashed symbols. It will easily be seen that the

\*It seems less important still when it is remembered that it is only found in the case of non-alternants, and that for the latter the results by the LCAO method are less reliable than for alternants because a self-consistent-field solution no longer results (11).

introduction of a non-zero overlap integral causes no change in the possibilities already discussed for the case where there are two residual molecules or radicals.

In addition to the method of calculating  $\mathcal{E}$  involving the direct solution of equations (8), which, as already mentioned, reduce to  $f(x) = 0$ , where  $f(x)$  is a polynomial of degree  $2n$ ,  $\mathcal{E}$  may be computed by a method due to Coulson(14).

He has shown that

$$\mathcal{E} = 2n\alpha - \frac{\gamma}{\pi} \int_{-\infty}^{\infty} \left\{ \frac{iyf'(iy)}{f(iy)} - 2n \right\} dy \quad \dots\dots\dots(23)$$

This equation only applies to alternants or non-alternants with no unoccupied energy levels having negative  $x_r$ , but it is only for such molecules that we are justified in computing  $B_{ab}$  from  $\mathcal{E}$  without necessarily considering the individual  $x_r$ . Details of the method of using (23) in numerical computations of  $\mathcal{E}$  are given by Coulson (loc. cit.).

It is also possible to calculate  $\mathcal{E}'$  by this method. The integral on the right-hand side of (23) is of course equal to  $2\pi$  times the sum of the negative roots of the equation  $f(x) = 0$  in accordance with equation (11). In an analogous way we may consider the quantities  $e_r$  to be the negative roots of another equation  $\eta(e) = 0$ .

By a similar line of reasoning to Coulson's(14, 7) it is easily deduced that if

$$\eta(e) = e^{2n} + A_1 e^{2n-1} + A_2 e^{2n-2} + \dots + A_{2n} \quad \dots\dots(24)$$

then we have the results

$$2 \sum_{r=1}^n e_r = -A_1 + \frac{1}{\pi} \int_{-\infty}^{\infty} \left\{ \frac{iy\eta'(iy)}{\eta(iy)} - 2n \right\} dy \quad \dots\dots\dots(25)$$

To employ (25) for computing  $\mathcal{E}'$  it is necessary to express  $\eta(e)$  and  $A_1$  in terms of  $f(x)$ . If it is assumed\* (cf. p. 568) that  $S = 0.25$  then it is fairly easily deduced that the required relationship between  $\eta(e)$  and  $f(x)$  is

\*In the more general case for an overlap integral  $S$ , equations (26) and (28) become

$$\eta(e) = \frac{(1 + Se)^{2n} f\left(\frac{e}{1 + Se}\right)}{\sum_{k=0}^{2n} B_k S^k} ; A_1 = \frac{\sum_{k=0}^{2n} B_k S^{k-1}}{\sum_{k=0}^{2n} B_k S^k}$$

$$\eta(e) = \frac{\left(\frac{4+e}{4}\right)^{2n} f\left(\frac{4e}{4+e}\right)}{\sum_{k=0}^{2n} B_k/4^k}, \quad \dots\dots\dots(26)$$

where  $B_k$  is the coefficient of  $x^{2n-k}$  in the equation

$$f(x) = x^{2n} + B_1x^{2n-1} + B_2x^{2n-2} + \dots + B_{2n} \quad \dots\dots\dots(27)$$

and finally by substituting (27) in (26) and collecting terms it can be shown that

$$A_1 = \frac{\sum_{k=1}^{2n} k B_k/4^{k-1}}{\sum_{k=0}^{2n} B_k/4^k} \quad \dots\dots\dots(28)$$

When using equation (25) in conjunction with (26) and (28) to calculate  $\xi'$  the same limitations with regard to non-alternants apply as for equation (23). It is easy to show that if (23) sums over occupied orbitals only, then the same is true for (25).

Another method of computing  $\xi$  depends upon the concept of conjugation energy which has been discussed in detail by Coulson and Longuet-Higgins(8). If two conjugated systems  $A$  and  $B$  are linked by a bond joining atoms  $a$  and  $b$ , these latter belonging to  $A$  and  $B$  respectively, then the resonance energy of the combined system exceeds the sum of the resonance energies of  $A$  and  $B$  by an amount  $C$  which is termed the conjugation energy. (Coulson and Longuet-Higgins used the symbol  $R$  for the conjugation energy as defined here, but it seems more convenient to reserve  $R$  for resonance energies). Thus the conjugation energy is defined by the equation

$$C = R - R_A - R_B, \quad \dots\dots\dots(29)$$

where  $R$ ,  $R_A$ ,  $R_B$ , are the resonance energies of the combined system, of  $A$ , and of  $B$ , respectively. Clearly equation (29) could alternatively be written

$$C = \xi_A + \xi_B - \xi \quad \dots\dots\dots(30)$$

The conjugation energy can be computed approximately from the table of conjugating powers given by Coulson and Longuet-Higgins(8), who showed that the mobile order,  $p_{ab}$ , of the bond linking the two systems, is a linear function of the geometrical mean of the conjugating powers of the atoms  $a$  and  $b$  in  $A$  and  $B$  respectively. From the graph given in their paper this empirical relationship is approximately expressed by

$$p_{ab} = 0.687 P_a P_b + 0.099, \quad \dots\dots\dots(31)$$

where  $P_a$ ,  $P_b$ , represent the square roots of the conjugating powers of  $a$  and  $b$  respectively, and using (31) in conjunction with the approximate result(8) that  $C = -p_{ab}\gamma$ ,



we obtain a useful equation for the approximate evaluation of  $C$

$$C = -(0.099 + 0.687P_aP_b)\gamma \quad \dots\dots\dots(32)$$

The result (32) is important because for many molecules the solution of (8) is tedious although the molecule is composed of two conjugated systems for each of which  $R$  is already known or is easily computed. It should be noted, however, that the empirical relation (31) was established with data for alternants only and it remains to test its validity for non-alternants. However, in the case of the latter the limitations already discussed for the use of resonance energies in the calculation of  $B$  again apply.

## V. RESULTS

The bond localization energies for a number of alternant hydrocarbons have been computed by the methods outlined in the preceding section. In Table 1 are listed the values of  $B_{ab}$ , calculated from equation (14) using  $x_r$ ,  $x_r^\dagger$ , computed by direct solution of the secular equations (8). The corresponding values of  $B'_{ab}$ , computed by means of equation (22), are also given, both in terms of  $-\beta$ , and in kcal. mol.<sup>-1</sup> using Dewar's value(10) of 34 kcal. mol.<sup>-1</sup> for  $-\beta$ , and the mobile orders,  $p_{ab}$ , calculated in the usual manner(2).

TABLE 1

Compound	Bond*	$p$	$B$ ( $-\gamma$ )	$B'$	
				( $-\beta$ )	kcal. mol. <sup>-1</sup>
Ethylene    ...    ...		1	0	0	0
Butadiene    ...    ...	1,2	0.894	0.472	0.175	6.0
	2,3	0.447	2.472	1.775	60
Hexatriene    ...    ...	1,2	0.871	0.516	0.212	7.2
	2,3	0.483	2.298	1.358	46
	3,4	0.785	0.988	0.387	13
Octatetraene    ...    ...	1,2	0.862	0.530	0.225	7.7
	2,3	0.495	2.053	1.395	47
	3,4	0.758	1.045	0.437	15
	4,5	0.529	1.861	1.233	42
Decapentaene    ...    ...	1,2	(0.869)	0.536	0.231	7.8
	2,3	(0.535)	1.999	1.341	46
	3,4	(0.780)	1.065	0.456	16
	4,5	(0.583)	1.761	1.136	39
	5,6	(0.766)	1.109	0.494	17
Benzene    ...    ...		0.667	1.528	0.892	30

TABLE 1—Continued

Compound	Bond*	$p$	$B$ ( $-\gamma$ )	$B'$	
				( $-\beta$ )	kcal. mol. <sup>-1</sup>
Naphthalene ... ..	1,2	0.725	1.259	0.653	22
	2,3	0.603	1.729	1.085	37
	1,9	0.555	2.166	1.251	43
	9,10	0.518	2.739	1.514	51
Anthracene ... ..	2,3	0.586	1.786	1.141	39
	11,12	0.485	2.888	1.655	56
Phenanthrene ... ..	2,3	0.623	1.66		
	9,10	0.775	1.065	0.484	16
	11,12	0.542	2.552	1.351	46
	12,13	0.461	2.856	1.669	56
Biphenyl ... ..	1,1'	0.370	3.455	2.217	75
	1,2	0.619	1.911	1.010	34
	2,3	0.677	1.443	0.829	28
	3,4	0.660	1.526	0.889	30
Styrene ... ..	$\alpha, \beta$	0.911	0.424	0.144	3.9
	1, $\alpha$	0.406	2.960	1.993	68
	1,2	0.610	1.952	1.036	35
	2,3	0.679	1.436	0.824	28
	3,4	0.659	1.525	0.889	30
Stilbene ... ..	$\alpha, \alpha'$	(0.83)	0.88	0.32	11
	1, $\alpha$	(0.45)	2.70	1.76	60
	1,2	(0.60)	1.98	1.06	36
1-Phenyl-1,3-butadiene	$\gamma, \delta$	(0.888)	0.516	0.212	7.2
	$\beta, \gamma$	(0.578)	1.556	1.199	41
	$\alpha, \beta$	(0.817)	0.941	0.356	12
	1, $\alpha$	(0.455)	2.648	1.715	58
	1,2	(0.598)	1.996	1.073	37
	2,3	(0.678)	1.423	0.810	28

\*The numbering system of *American Chemical Abstracts*, with obvious extensions, has been employed here.

To facilitate the calculation of  $B$  and  $B'$  for other bonds the resonance energies,  $R$ ,  $R'$ , of various molecules and fragments have been collected in Table 2. All the values listed have been computed directly using equations (13) or (21) to at least one significant figure further than the tabulated value. It should be borne in mind that the resonance energies of the radicals listed in Table 2 have been defined differently (pp. 565-6) from the resonance energies of free radicals as defined for example by Wheland(4).

The results presented in Table I show clearly that there is a close correspondence between  $p$  and  $B$  and  $B'$ ; in fact the greater the value of  $p$  the smaller is  $B$  or  $B'$ . It would be expected that, other things being equal, for the bonds of any given

TABLE 2

Compound	$R$ ( $-\gamma$ )	$R'$ ( $-\beta$ )
Ethylene ... ..	0	0
Butadiene ... ..	0.472 <sup>a</sup>	0.175
Hexatriene ... ..	0.988 <sup>b</sup>	0.387
Octatetraene ... ..	1.518 <sup>a</sup>	0.612
Decapentaene ... ..	2.053	0.843
2-Vinyl-butadiene ... ..	0.899 <sup>c</sup>	0.322
Benzene ... ..	2 <sup>a</sup>	1.067 <sup>d</sup>
Naphthalene ... ..	3.683 <sup>a</sup>	1.863 <sup>d</sup>
Anthracene ... ..	5.314 <sup>a</sup>	2.608 <sup>d</sup>
Phenanthrene ... ..	5.448 <sup>a</sup>	2.736 <sup>d</sup>
Biphenyl ... ..	4.383 <sup>a</sup>	2.252 <sup>d</sup>
Styrene ... ..	2.424 <sup>d</sup>	1.211 <sup>d</sup>
1-Phenyl-butadiene ... ..	2.941 <sup>e</sup>	1.422
2-Phenyl-butadiene ... ..	2.857 <sup>f</sup>	1.362
<i>o</i> -Benzoquinodimethane ... ..	1.954 <sup>g</sup>	0.778
Carbon ... ..	-1	-0.8
Allyl ... ..	-0.172	-0.310
Pentadienyl ... ..	0.464	0.017
Heptatrienyl ... ..	1.055	0.302
Nonatetraenyl ... ..	1.628	0.570
Benzyl ... ..	1.721	0.671
Cinnamyl ... ..	2.385	1.023

<sup>a</sup> In agreement with the value given by Hückel(3) to two decimals.

<sup>b</sup> The value  $-0.96\gamma$  given by Hückel(3) is in error.

<sup>c</sup> In agreement with the value  $-0.90\gamma$  given by Syrkin and Diatkina (15).

<sup>d</sup> These values agree with those given by Wheland(4) to two decimal places.

<sup>e</sup> Syrkin and Diatkina(15) give the incorrect value  $-2.90\gamma$ .

<sup>f</sup> Syrkin and Diatkina(15) give the incorrect value  $-2.78\gamma$ .

<sup>g</sup> Diatkina and Syrkin(17) give the incorrect value  $-1.92\gamma$ . Pullman, Berthier, and Pullman(18) give  $-1.956\gamma$ , and from the roots listed by them it is apparent that this small difference from the present value is due to cumulative errors in the third decimal. The present value was computed retaining five decimals throughout the calculation.

molecule the one closest to a pure double bond (i.e. of highest mobile order) will require the least energy to perturb it to a double bond; but the relationship goes further than this. It is found that the same applies to the comparison of bonds in *different* molecules. This latter result means that  $B$  and  $B'$  depend almost solely upon  $p$  and are approximately independent of the remainder of the molecule. Some exceptions occur, especially for bonds of low mobile order (below 0.6). Figures

1 and 2 show graphically the relations between  $p$  and  $B$  and between  $p$  and  $B'$  respectively. They show 20 points for the bonds of the compounds ethylene, butadiene, hexatriene, octatetraene, benzene, styrene, and naphthalene; the lines in these figures were drawn by inspection rather than by a "least squares" or similar method.

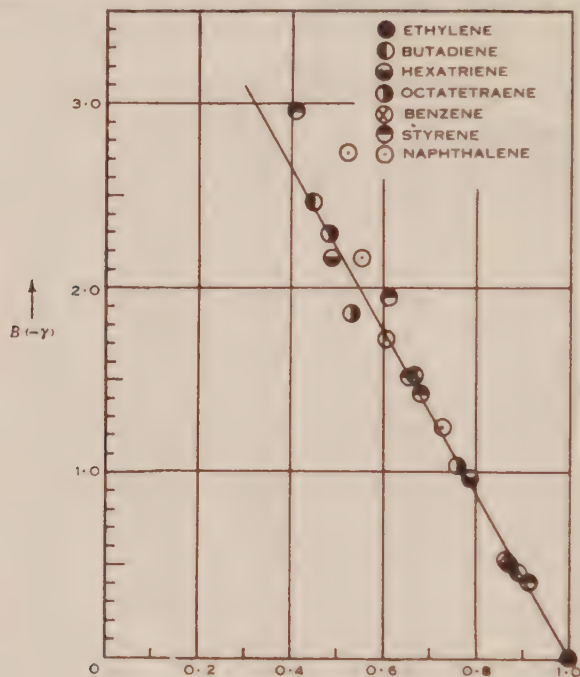


Fig. 1

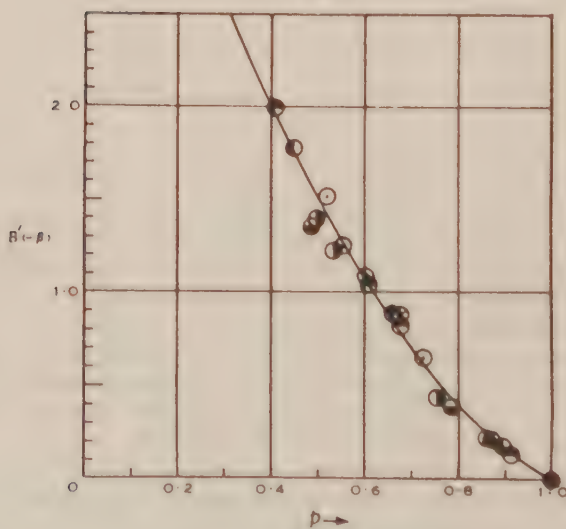


Fig. 2

butadiene, hexatriene, octatetraene, benzene, styrene, and naphthalene; the lines in these figures were drawn by inspection rather than by a "least squares" or similar method.



For some purposes it is convenient to have analytical expressions for these relationships. The equations to the curves drawn in Figures 1 and 2 respectively are

$$p = 1 + 0.226B/\gamma, \quad \dots\dots\dots(33)$$

$$p = 1.2 - (0.04 - 0.3B'/\beta)^{1/2} \quad \dots\dots\dots(34)$$

The values of  $p$  in Table 1 are given in parenthesis when the directly computed values

TABLE 3

Compound	Bond	$p_B$	$d$	$p_{B'}$	$d'$	$p_{cal.}$
Ethylene ...		1	0	1	0	1
Butadiene ...	1,2	0.89	0	0.90	-0.01	0.89
	2,3	0.44	+0.01	0.44	+0.01	0.45
Hexatriene ...	1,2	0.88	-0.01	0.89	-0.02	0.87
	2,3	0.48	0	0.53	-0.05	0.48
	3,4	0.78	+0.01	0.81	-0.02	0.79
Octatetraene ...	1,2	0.88	-0.02	0.87	-0.01	0.86
	2,3	0.54	-0.04	0.52	-0.02	0.50
	3,4	0.76	0	0.79	-0.03	0.76
	4,5	0.58	-0.05	0.56	-0.03	0.53
Benzene ...		0.66	+0.01	0.65	+0.02	0.67
Naphthalene ...	1,2	0.72	+0.01	0.71	+0.02	0.73
	2,3	0.61	-0.01	0.60	0	0.60
	9,1	0.51	+0.05	0.56	0	0.56
	9,10	0.38	+0.14	0.50	+0.02	0.52
Anthracene ...	2,3	0.60	-0.01	0.58	+0.01	0.59
	11,12	0.35	+0.14	0.47	+0.02	0.49
Phenanthrene ...	9,10	0.76	+0.02	0.77	+0.01	0.78
	11,12	0.42	+0.12	0.53	+0.01	0.54
Biphenyl ...	1,1'	0.22	+0.15	0.36	+0.01	0.37
	1,2	0.57	+0.05	0.61	+0.01	0.62
	2,3	0.67	+0.01	0.66	+0.02	0.68
	3,4	0.66	0	0.65	+0.01	0.66
Styrene ...	$\alpha, \beta$	0.90	+0.01	0.91	0	0.91
	1, $\alpha$	0.33	+0.08	0.40	+0.01	0.41
	1,2	0.56	+0.05	0.61	0	0.61
	2,3	0.68	0	0.66	+0.02	0.68
	3,4	0.66	0	0.65	+0.01	0.66

have not been derived and are not recorded in the literature. These values have been calculated by means of equation (34).

The reliability of mobile orders derived from equations (33) and (34) may be judged from comparison with values of  $p$  obtained by direct calculation of  $2 \sum c_j c_k$  as given in Table 3. The columns headed  $d$ ,  $d'$ , represent  $p_{\text{cal.}} - p_B$ ,  $p_{\text{cal.}} - p_{B'}$ , respectively. An analysis of the data of Table 3, excluding the figures for ethylene, shows that using  $B$ , the standard deviation,  $\sigma$ , in computing  $p$  is 0.059, while using  $B'$  it is 0.017, indicating that (34) is the more reliable relationship. However, inspection of Figure 1 or of Table 3 makes it apparent that results obtained by equation (33) are considerably better when the right-hand side is more than 0.6, and an analysis of the appropriate data shows that in the latter event  $\sigma$  is 0.009; for (34) in the same range  $\sigma$  is roughly the same as over the whole range. Thus (33) will be very accurate for reactive bonds.

TABLE 4

Compound	Bond	$p$	$B$ ( $-\gamma$ )
Anthracene ...	1,2	0.738	1.22
	13,1	0.535	2.27
	9,13	0.606	1.87
Phenanthrene ...	1,2	0.705	1.34
	3,4	0.702	1.36
	11,1	0.575	2.09
	10,11	0.506	2.40
	4,12	0.590	2.00
Stilbene ...	2,3	—	1.43
	3,4	—	1.56
1-Phenyl-butadiene	3,4	—	1.55

Inspection of Table 1 shows that for some of the molecules listed there, e.g. anthracene, phenanthrene, stilbene, no data are provided for some of the bonds. For these bonds the resonance energies of the residual molecules are rather troublesome to compute directly due to lack of symmetry of the molecule. However, an approximate calculation using equations (29) and (32) and the table of  $P^2$  given by Coulson and Longuet-Higgins(8) is easily carried through. Values of  $B$  obtained in this way are recorded in Table 4 and the relevant resonance energies and other data are given in Table 5. It may be noted that in the calculation of  $R$  for 1-phenyl-3-vinyl-butadiene it was assumed that the two butadiene substituents have independent additive effects upon  $R$ , i.e. that  $P_3$  for 1-phenyl-butadiene is approximately equal to  $P_3$  for butadiene itself. This seems justifiable because  $P_3$  for butadiene is 0.63 and for 1-vinyl-butadiene (hexatriene) it is 0.62; in addition, as  $P$  is greater for vinyl than phenyl, the latter should have the smaller effect upon the  $P_3$  of the butadiene part of the molecule. The resonance energy thus obtained is less than

that for 1-phenyl-hexatriene, in agreement with expectation, phenyl-vinyl-butadiene being the more branched isomer.

TABLE 5

Compound	<i>C</i> ( $-\gamma$ )	<i>R</i> ( $-\gamma$ )
$\alpha$ -Vinyl-naphthalene ... ..	0.423	4.11
$\beta$ -Vinyl-naphthalene ... ..	0.407	4.09
1-Phenyl-hexatriene ... ..	0.458	3.45
2-Vinyl-hexatriene ... ..	0.402	1.39
1-Phenyl-3-vinyl-butadiene ...	0.849	3.32
1-( <i>o</i> -Vinyl-phenyl)butadiene ...	0.458	3.35
Butadienyl- <i>o</i> -quinodimethane	0.63	3.05

The data for butadienyl-*o*-quinodimethane and for 1-(*o*-vinyl-phenyl)butadiene are dependent upon the values of *P* for the methylene carbon of *o*-quinodimethane and for the *ortho*-carbon of styrene respectively. These values of *P* were computed from the data of Syrkin and Diatkina(15) for vinyl-*o*-quinodimethane, and of Berthier and Pullman(16) for *o*-divinylbenzene; the results were respectively 0.97 and 0.66.

The relationships connecting *B* and *B'* with bond lengths, the bond localization energies for long polyenes, for non-alternants, and for heterocyclic systems, and the application of bond localization energy to problems of molecular structure and reactivity will be considered in subsequent papers.

## VI. ACKNOWLEDGMENT

I should like to thank Mr. A. N. Hambly for his constructive criticisms during the preparation of the manuscript.

## VII. REFERENCES

- (1) WHELAND, G. W., and PAULING, L.—*J. Amer. Chem. Soc.* **57**: 2091 (1935).
- (2) COULSON, C. A.—*Proc. Roy. Soc. A* **169**: 413 (1939).
- (3) WHELAND, G. W.—*J. Chem. Phys.* **2**: 474 (1934); Hückel, E.—*Z. Elektrochem.* **43**: 752, 827 (1937).
- (4) WHELAND, G. W.—*J. Amer. Chem. Soc.* **63**: 2025 (1941).
- (5) COULSON, C. A.—*Discussions Faraday Soc.* **2**: 16 (1947).
- (6) WHELAND, G. W.—*J. Amer. Chem. Soc.* **64**: 900 (1942).
- (7) COULSON, C. A., and LONGUET-HIGGINS, H. C.—*Proc. Roy. Soc. A* **191**: 39 (1947).
- (8) COULSON, C. A.—*Ibid.* **195**: 188 (1948).
- (9) KOOYMAN, E. C., and KETELAAR, J. A. A.—*Rec. Trav. Chim.* **65**: 859 (1946).
- (10) DEWAR, M. J. S.—*Trans. Faraday Soc.* **42**: 767 (1946).
- (11) COULSON, C. A., and RUSHBROOKE, G. S.—*Proc. Camb. Phil. Soc.* **36**: 193 (1940).

- (12) MULLIKEN, R. S., RIEKE, C. A., and BROWN, W. G.—*J. Amer. Chem. Soc.* **63** : 48 (1941).
- (13) MULLIKEN, R. S., and RIEKE, C. A.—*Ibid.* **63** : 1770 (1941).
- (14) COULSON, C. A.—*Proc. Camb. Phil. Soc.* **36** : 201 (1940).
- (15) SYRKIN, Y. K., and DIATKINA, M. E.—*Acta Physicochim.* **21** : 641 (1946).
- (16) BERTHIER, G., and PULLMAN, B.—*Bull. Soc. Chim. Fr.* **1948** : 554 (1948).
- (17) DIATKINA, M. E., and SYRKIN, Y. K.—*Acta Physicochim.* **21** : 921 (1946).
- (18) PULLMAN, A., BERTHIER, G., and PULLMAN, B.—*Bull. Soc. Chim. Fr.* **1948** : 450 (1948).



# THE RELATIVE STABILITY OF INTERNAL METAL COMPLEXES

## II. METAL DERIVATIVES OF 8-HYDROXYQUINOLINE 5-SULPHONIC ACID AND A SERIES OF MONOCARBOXYLIC MONO- $\alpha$ -AMINO ACIDS INCLUDING HISTIDINE

By L. E. MALEY\* and D. P. MELLOR\*

[*Manuscript received May 23, 1949*]

### *Summary*

Stability constants of a series of metal complexes derived from the series of  $\alpha$ -amino acids, glycine, alanine, leucine, valine, histidine, and 8-hydroxyquinoline 5-sulphonic acid have been determined by an electrometric method.

The metals fall in the same order, as regards stability of their complexes, as was found for the chelating molecules studied in Part I of this series.

As the length of the  $\alpha$  chain of the amino acid increases, there is a small decrease in the stability of the amino acid complexes of any one metal. The cobalt II complex of histidine is much more stable than cobalt II complexes derived from any of the other  $\alpha$ -amino acids. It is suggested that this explains why histidine is able to nullify the inhibiting action of cobalt on the growth and respiration of various aerobic and anaerobic organisms.

Though there is no great difference in the stability of corresponding metal complexes derived from 8-hydroxyquinoline and 8-hydroxyquinoline 5-sulphonic acid (when allowance is made for different solvents used in making the measurements), there is a marked difference in the bacteriostatic action of these two chelating molecules. A reason for this difference is suggested.

The stability constants of any chelating reaction involving the liberation of any number of hydrogen ions can be evaluated by using the generalized mathematical treatment which has been set out in this series of papers.

### I. INTRODUCTION

In Part I of this series, results were reported on the measurements of stability constants of metal complexes of salicylaldehyde, 8-hydroxyquinoline, and acetylacetone. This paper deals with chelating molecules of a different kind, namely those which in solution exist as zwitterions. All the metal complexes derived from the zwitterions proved sufficiently soluble to be studied in water solutions. For the greater part, the chelating molecules were chosen because of the possible biochemical interest of the stability constants of their metal complexes. Thus the amino acids studied were the first four of the series of monocarboxylic mono- $\alpha$ -amino acids which are products of the hydrolysis of proteins(1). Histidine was added to the list because it has been found(2) that some metals, especially cobalt, are effective inhibitors of the growth and respiration of various microorganisms in animal tissues and tumors and that among the amino acids that have been tested(3, 4), histidine alone can

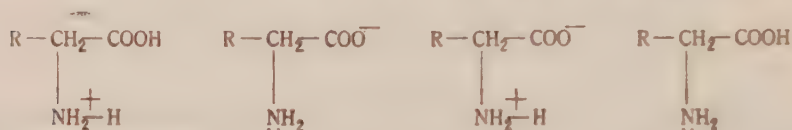
\*Chemistry Department, University of Sydney.

nullify the inhibiting action of cobalt. 8-Hydroxyquinoline 5-sulphonic acid was studied because of the interest in comparing the stability of its water soluble metal complexes with that of the corresponding insoluble complexes of 8-hydroxyquinoline itself. Added interest lies in the fact that 8-hydroxyquinoline 5-sulphonic acid is, unlike the unsubstituted molecule, not bacteriostatic.

## II. THEORETICAL

### (a) Chelation of $\alpha$ -Amino Acids

There are four forms in which the  $\alpha$ -amino acids may be written, namely,

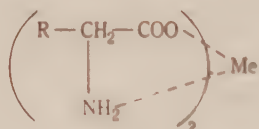


These may conveniently be denoted by the symbols  $\text{HAH}^+$ ,  $\text{A}^-$ ,  $\text{A}^-\text{H}^+$ , and  $\text{HA}$  respectively\*.

The dissociation constants relevant to the calculation of stability constants are

$$k_{\text{NH}_2} = \frac{(\text{H}^+)(\text{A}^-)}{(\text{A}^-\text{H}^+)} \text{ and } k_{\text{COOH}} = \frac{(\text{H}^+)(\text{A}^-\text{H}^+)}{(\text{HAH}^+)}.$$

Complexes formed with bivalent metals have the general formula



which will be abbreviated to  $\text{MeA}_2$ .

For the definition of the symbols used in the argument below see Part I (6):

$$\text{T}_{\text{Me}} = (\text{Me}^{++}) + (\text{Me}^+\text{A}) + (\text{MeA}_2)$$

$$\text{T}_{\text{L}} = (\text{A}^-) + (\text{A}^-\text{H}^+) + (\text{HAH}^+) + \bar{n} \cdot \text{T}_{\text{Me}}$$

$$\text{T}_{\text{H}} = (\text{H}^+) + (\text{A}^-\text{H}^+) + 2(\text{HAH}^+)$$

$$n_{\text{A}} = \frac{(\text{A}^-\text{H}^+) + 2(\text{HAH}^+)}{(\text{A}^-) + (\text{A}^-\text{H}^+) + (\text{HAH}^+)}.$$

Eliminating  $(\text{A}^-) + (\text{A}^-\text{H}^+) + (\text{HAH}^+)$  from  $n_{\text{A}}$  and  $\text{T}_{\text{L}}$  and substituting for  $\text{T}_{\text{H}}$

$$\text{T}_{\text{L}} = \frac{\text{T}_{\text{H}} - (\text{H}^+)}{n_{\text{A}}} + \bar{n} \cdot \text{T}_{\text{Me}}.$$

The chelate molecule was added to the reaction solution with one dissociable

\*Since  $\alpha$ -amino acids in solution exist as zwitterions (5) the concentration of form  $\text{HA}$  will be negligible.

hydrogen, therefore an equivalent expression for  $T_H$  is

$$T_H = (E - Na) + \frac{K_w}{(H^+)} + T_L$$

$$\bar{n} = \frac{T_L \left(1 - \frac{1}{\bar{n}_A}\right) - \left[ \frac{(E - Na) + \left(\frac{K_w}{H^+}\right) - (H^+)}{\bar{n}_A} \right]}{T_{Me}}$$

where

$$\bar{n}_A = \frac{1 + 2 \frac{(H^+)}{k_{COOH}}}{1 + \frac{k_{NH_2}}{(H^+)} + \frac{(H^+)}{k_{COOH}}}$$

*Derivation of an Expression for  $\text{Log } \frac{1}{(A^-)}$*

$$\text{Log } \frac{1}{(A^-)} = \log (H^+) - \log k_{NH_2} - \log (A^-H^+)$$

Total uncombined ligand =  $T_U = (A^-) + (HAH^+) + (A^-H^+)$ .

$\text{Log } (A^-H^+) = \log (T_L - \bar{n}T_{Me})$  (1 - fraction of  $T_U$  as dissociated acid - fraction of  $T_U$  as undissociated base). The percentages of dissociated acid and undissociated base were obtained by using Henderson's equations as outlined in Part I (6).

The acid dissociation constants of the amino acids used in the calculations were determined potentiometrically by titration with standard alkali and acid and the following  $pK_a$  values were obtained in water at  $25^\circ \pm 0.02^\circ\text{C}$ . :

	$pK_a(NH_2)$	$pK_a(COOH)$
Glycine	9.78	2.35
$\alpha$ -Alanine	9.87	2.35
Valine*	9.62	2.32
Leucine*	9.60	2.36

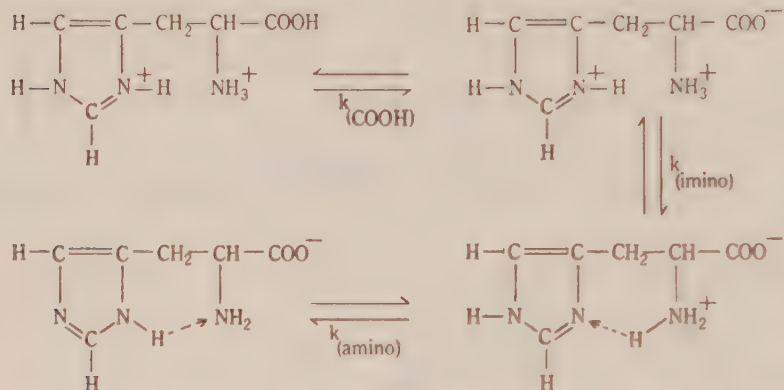
These values agree with those quoted in the literature (7).

\*If the dissociation constants of valine and leucine are measured at infinite dilution at  $25^\circ\text{C}$ . the corresponding values (9) are

	$pK_a(NH_2)$	$pK_a(COOH)$
Valine	9.72	2.29
Leucine	9.74	2.33

(b) *The Chelation of Histidine*

Histidine can exist in solution in four different forms (8) as follows :



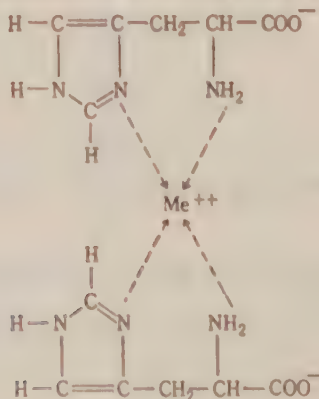
These will be denoted by  $\text{H.AH}_2^{++}$ ,  $\text{AH}_2^{++}$ ,  $\text{AH}_1^+$ , and  $\text{A}$  respectively.

In attempting to decide which groups in histidine participate in the formation of chelate rings with metal atoms it is necessary to keep the following experimental facts in mind.

(i) After chelation, the amino group is no longer titratable. For this reason it would appear that the amino group is involved in chelation.

(ii) Notwithstanding the fact that the values of  $\text{pK}_{\text{NH}_2}$  and  $\text{pK}_{\text{COOH}}$  are substantially the same for both the simple amino acids and histidine, the stability of the histidine complexes proved to be greater than that of the corresponding metal complexes of the simple amino acids (see chelation titration curves Figs. 1 and 2). This suggests that the mode of chelation of histidine differs from that of the other amino acids. As the amino group is involved, the only other group that could be concerned is the imino group.

The fact that the coordination of cobaltous ions with histidine occurs at physiologically significant and lower pH values than the other  $\alpha$ -amino acids\* is further evidence supporting a formulation of the complex as follows :



\*This has been suggested (3) as the basis of the biological specificity of histidine.



The three acid dissociation constants of histidine which must be taken into account in considering its chelation are

$$k_{\text{NH}_2} = \frac{(\text{A}^-)(\text{H}^+)}{(\text{A}^-\text{H}_1^+)}$$

$$k_{\text{NH}} = \frac{(\text{A}^-\text{H}_1^+)(\text{H}^+)}{(\text{A}^-\text{H}_2^{++})}$$

$$k_{\text{COOH}} = \frac{(\text{A}^-\text{H}_2^{++})(\text{H}^+)}{(\text{HAH}_2^{++})}$$

Expressions for  $\bar{n}$  and  $\log \frac{1}{(\text{A}^-)}$  are derived by using the following equations:

$$\text{T}_L = (\text{A}^-) + (\text{A}^-\text{H}^+) + (\text{A}^-\text{H}^{++}) + (\text{HAH}_2^{++}) + \bar{n} \cdot \text{T}_{\text{Me}}$$

$$\text{T}_H = (\text{H}^+) + (\text{A}^-\text{H}_1^+) + 2(\text{A}^-\text{H}_2^{++}) + 3(\text{HAH}_2^{++})$$

$$\bar{n}_A = \frac{(\text{A}^-\text{H}_1^+) + 2(\text{A}^-\text{H}_2^{++}) + 3(\text{HAH}_2^{++})}{(\text{A}^-) + (\text{A}^-\text{H}_1^+) + (\text{A}^-\text{H}_2^{++}) + (\text{HAH}_2^{++})}$$

$$\begin{aligned} &= \frac{k_{\text{NH}} \cdot (\text{H}^+) + 2(\text{H}^+)^2 + 3 \left[ \frac{(\text{H}^+)^3}{k_{\text{COOH}}} \right]}{k_{\text{NH}_2} \cdot k_{\text{NH}} + k_{\text{NH}} \cdot (\text{H}^+) + (\text{H}^+)^2 + \left[ \frac{(\text{H}^+)^3}{k_{\text{COOH}}} \right]} \\ \alpha &= \frac{(\text{A}^-)}{(\text{A}^-) + (\text{A}^-\text{H}_1^+) + (\text{A}^-\text{H}_2^{++}) + (\text{HAH}_2^{++})} \\ &= \frac{k_{\text{NH}_2} \cdot k_{\text{NH}}}{k_{\text{NH}_2} \cdot k_{\text{NH}} + k_{\text{NH}} \cdot (\text{H}^+) + (\text{H}^+)^2 + \left[ \frac{(\text{H}^+)^3}{k_{\text{COOH}}} \right]} \end{aligned}$$

Eliminating  $(\text{A}^-) + (\text{A}^-\text{H}_1^+) + (\text{A}^-\text{H}_2^{++}) + (\text{HAH}_2^{++})$  from the equations for  $\text{T}_L$  and  $\bar{n}_A$  and substituting for  $\text{T}_H$

$$\bar{n} = \frac{\text{T}_L}{\text{T}_{\text{Me}}} - \left[ \frac{\text{T}_H - (\text{H}^+)}{\bar{n}_A \cdot \text{T}_{\text{Me}}} \right]$$

Since histidine was added to the reaction solution with two dissociable hydrogens,

$$\text{the total acid concentration} = \text{T}_H = (\text{E} - \text{Na}) + \frac{K_w}{(\text{H}^+)} + 2\text{T}_L,$$

therefore

$$\bar{n} = \frac{\text{T}_L}{\text{T}_{\text{Me}}} \left( 1 - \frac{2}{\bar{n}_A} \right) - \left[ \frac{(\text{E} - \text{Na}) + \frac{K_w}{(\text{H}^+)} - (\text{H}^+)}{\bar{n}_A \cdot \text{T}_{\text{Me}}} \right]$$

Instead of using Henderson's equations for deriving values of  $\log \frac{1}{(\text{A}^-)}$  for

histidine, the alternative method referred to in Part I was applied.

Thus combining equations  $T_L$  and  $\alpha$  and substituting for  $T_H$

$$(A^-) = \frac{\alpha}{n_A} (T_H - (H^+)),$$

$$\log \frac{1}{(A^-)} = \log \left\{ \frac{k_{NH_2} \cdot k_{NH} + k_{NH} \cdot (H^+) + (H^+)^2 + \left[ \frac{(H^+)^2}{k_{COOH}} \right]}{k_{NH_2} \cdot k_{NH}} \right\} \left\{ \frac{\bar{n}_A}{T_H - (H^+)} \right\}$$

where

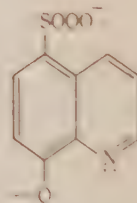
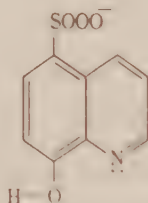
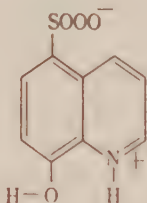
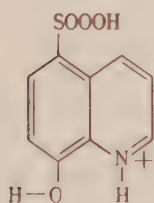
$$T_H = \left[ (E - Na) + \frac{K_w}{(H^+)} + 2T_L \right].$$

The acid dissociation constants of histidine used in the calculations were determined potentiometrically by titration with standard alkali and acid. The following results were obtained in water at  $25^\circ \pm 0.02^\circ \text{C}$ .:  $pK_a(\text{COOH}) = 1.77$ ,  $pK_a(\text{imino}) = 6.10$ , and  $pK_a(\text{amino}) = 9.18$ .

These results agree with the results reported (7, 9) in recent publications.

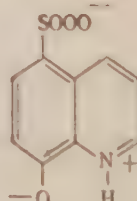
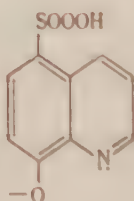
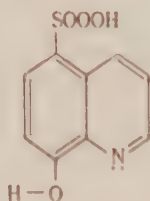
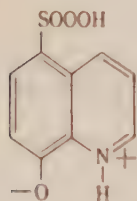
(c) *Chelation of 8-Hydroxyquinoline 5-Sulphonic Acid\**

There are four possible molecular forms of 8-hydroxyquinoline-5-sulphonic acid in solution, namely,



These can be conveniently denoted by the symbols  $H_2AH^+$ ,  $HA \cdot H^+$ ,  $HA$ , and  $A^-$  respectively.

Owing to the fact that 8-hydroxyquinoline 5-sulphonic acid is a zwitterion forming molecule and that  $k_{OH} \ll k_{NH}$ , the following ionic forms



may be neglected.

\*Although complexes of the type  $[MeA_4]^-$  are not internal complexes the same principles as have been applied to internal complexes may be used for deriving expressions for determining their stability.

The three hydrogen ion dissociation constants are

$$k_{\text{SO}_3\text{H}} = \frac{(\text{H}^+) (\text{HA}^-\text{H}^+)}{(\text{H}_2\text{AH}^+)}$$

$$k_{\text{NH}} = \frac{(\text{H}^+) (\text{HA}^-)}{(\text{HA}^-\text{H}^+)}$$

$$k_{\text{OH}} = \frac{(\text{H}^+) (\text{A}^-)}{(\text{HA}^-)}$$

$$\bar{n} = \frac{\text{attached ligand}}{\text{T}_{\text{Me}}}.$$

$$\text{T}_{\text{L}} = (\text{A}^-) + (\text{HA}^-) + (\text{HA}^-\text{H}^+) + (\text{H}_2\text{AH}^+) + \bar{n} \cdot (\text{T}_{\text{Me}})$$

$$\text{T}_{\text{H}} = (\text{H}^+) + (\text{HA}^-) + 2(\text{HA}^-\text{H}^+) + 3(\text{H}_2\text{AH}^+)$$

$$\bar{n}_{\text{A}} = \frac{(\text{HA}^-) + 2(\text{HA}^-\text{H}^+) + 3(\text{H}_2\text{AH}^+)}{(\text{A}^-) + (\text{HA}^-) + (\text{HA}^-\text{H}^+) + (\text{H}_2\text{AH}^+)}$$

$$= \frac{1 + 2 \frac{(\text{H}^+)}{k_{\text{NH}}} + 3 \left[ \frac{(\text{H}^+)}{k_{\text{NH}}} \cdot \frac{(\text{H}^+)}{k_{\text{SO}_3\text{H}}} \right]}{1 + \frac{k_{\text{OH}}}{(\text{H}^+)} + \frac{(\text{H}^+)}{k_{\text{NH}}} + \frac{(\text{H}^+)}{k_{\text{NH}}} \cdot \frac{(\text{H}^+)}{k_{\text{SO}_3\text{H}}}}$$

Since the chelating molecule, 8-hydroxyquinoline 5-sulphonic acid, was added to the reaction solution with two dissociable hydrogens, the expressions for  $\text{T}_{\text{H}}$  and  $\bar{n}$  will be formally identical with those for histidine hydrochloride.

The expression for  $\log \frac{1}{(\text{A}^-)}$  is derived as follows:

$$\text{T}_{\text{U}} = (\text{A}^-) + (\text{HA}^-) + (\text{HA}^-\text{H}^+) + (\text{H}_2\text{AH}^+).$$

Therefore the concentration of  $\text{HA}^-$  equals the concentration of the  $\text{T}_{\text{U}}$  molecules which have the following attached groups simultaneously present: a dissociated  $\text{SO}_3\text{H}$  group, a dissociated  $\text{NH}$  group, and an undissociated  $\text{OH}$  group.

$$\text{Further } \text{T}_{\text{U}} = \text{T}_{\text{L}} - \bar{n} \cdot \text{T}_{\text{Me}},$$

therefore

$$\log (\text{HA}^-) = \log (\text{T}_{\text{L}} - \bar{n} \cdot \text{T}_{\text{Me}}) \left\{ \begin{array}{l} \text{fraction of } \text{T}_{\text{U}} \text{ having a} \\ \text{dissociated } \text{SO}_3\text{H} \text{ group} \end{array} \right\}$$

$$\times \left\{ \begin{array}{l} \text{fraction of } \text{T}_{\text{U}} \text{ having a} \\ \text{dissociated } \text{NH} \text{ group} \end{array} \right\} \times \left\{ \begin{array}{l} \text{fraction of } \text{T}_{\text{U}} \text{ having an} \\ \text{undissociated } \text{OH} \text{ group} \end{array} \right\}$$

$$\log \frac{1}{(\text{A}^-)} = \log (\text{H}^+) - \log k_{\text{OH}} - \log (\text{HA}^-).$$

Henderson's equations were used to give the respective percentages of molecules with dissociated and undissociated groups.

The acid dissociation constants of 8-hydroxyquinoline 5-sulphonic acid used

in the calculations were determined potentiometrically by titration with standard alkali and acid. The  $pK_a$  values obtained in water at 25 °C. were  $pK_a(\text{SO}_3\text{H}) = 2.04$ ,  $pK_a(\text{NH}) = 4.19$ , and  $pK_a(\text{OH}) = 8.48$ .

These values agree with those reported by Albert(10).

### III. EXPERIMENTAL

#### (a) $\alpha$ -Amino Acids: *Glycine, Alanine, Leucine, and Valine*

The solutions which were titrated potentiometrically with 1.0434N NaOH contained 0.001M metal perchlorate,\* 0.005M  $\alpha$ -amino acid and sufficient free perchloric acid to give a pH of approximately 1.7.

Additions of NaOH were made by means of a microburette and the glass electrode was standardized as outlined in Part I. All titrations were made in duplicate; the pH values resulting from the addition of the same amount of NaOH in two consecutive titrations were reproducible to within  $\pm 0.01$  of a pH unit. Free perchloric acid was determined by a separate titration on the same solution but with the  $\alpha$ -amino acid omitted. Titrations of cobalt and manganese were carried out in an oxygen-free nitrogen atmosphere.

With some metals the chelation reaction could be roughly followed by colour changes. Thus with copper the solution was practically colourless at the beginning of the titration but as it progressed, a bright blue colour gradually developed reaching its maximum at the end of the pH range over which chelation occurs. C.P.  $\alpha$ -amino acids were used throughout.

#### (b) *Histidine*

The same conditions were used as for the other  $\alpha$ -amino acids except that the concentration of histidine was 0.002M. As other workers have noted, the cobalt complex was found to be very sensitive to oxygen, on exposure to which it rapidly turns an amber brown†. At the end of a titration under the conditions employed the solution was quite colourless indicating that it was oxygen free.

A.R. histidine monohydrochloride was used throughout this work.

#### (c) *8-Hydroxyquinoline 5-Sulphonic Acid*

The conditions used were the same as those for histidine (0.001M metal perchlorate and 0.002M 8-hydroxyquinoline 5-sulphonic acid). The 8-hydroxyquinoline 5-sulphonic acid used melted at 321°C.

### IV. RESULTS

#### (a) $\alpha$ -Amino Acids

A typical set of titration curves obtained with the chelation of  $\alpha$ -amino acids is

\*For the palladium estimation 0.001M  $\text{K}_2\text{PdCl}_4$  was used instead of the perchlorate.

†Histidine cobalt complex rapidly absorbs oxygen from the air in two stages (2) forming in both stages coloured compounds.



shown in Figure 1 and a summary of the stability constants is given in Table 2,

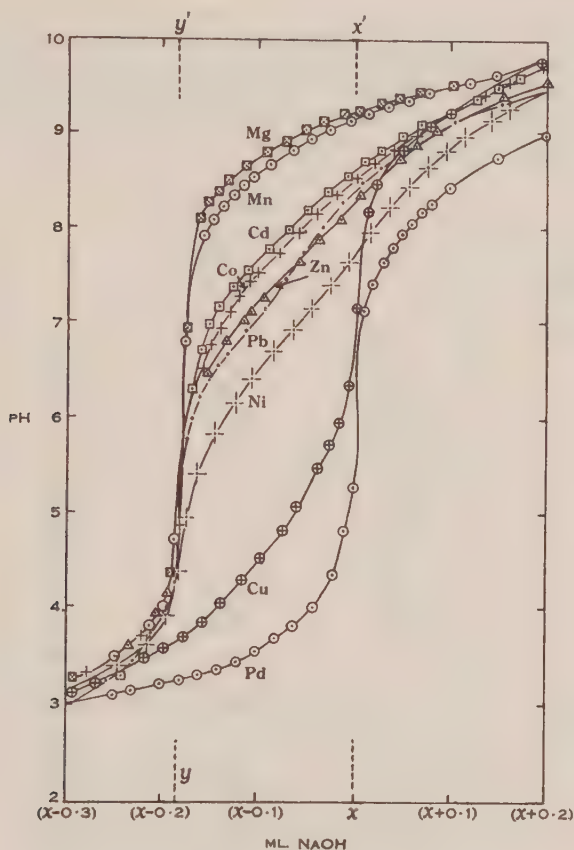


Fig. 1.—The chelation titration curves of metal-glycine complexes.

The line ( $yy'$ ) represents the excess acid present in the solution.

The line ( $xx'$ ) represents titration of excess acid plus the two moles of hydrogen ion liberated during chelation.

The excess acid in the different titrations was :

With Pd = 2.371 ml. of 1.0885N NaOH

Cu = 3.956 " " " "

Ni = 3.505 " " " "

Pb = 3.31 " " " "

Zn = 2.815 " " " "

Co = 3.149 " " " "

Cd = 3.473 " " " "

Mn = 3.473 ml. of 1.0434N NaOH

Mg = 3.215 ml. of 1.0885N NaOH.

#### (b) Histidine

The chelation titration curves obtained with histidine are shown in Figure 2 and the corresponding formation curves in Figure 3. The formation curve data

are summarized in Table 1 and the stability constants are set out in Table 3.

TABLE 1  
FORMATION CURVE DATA FOR THE HISTIDINE COMPLEXES AT 25°C.  
Concentration of Metal = 0.001M; Histidine Monohydrochloride = 0.002M  
Initial Volume 100 ml.

Metal	T <sub>H</sub> in ml. of 1.0434N Acid	pH	Titration of NaOH in ml. of 1.0434N NaOH	$\bar{n}$	$\text{Log } \frac{1}{(A^-)}$
Copper	3.68	3.70	3.37	0.46	10.71
		3.78	3.38	0.50	10.56
		3.86	3.40	0.62	10.43
		5.40	3.596	1.496	7.87
		5.48	3.60	1.502	7.74
Cobalt	3.57	5.23	3.29	0.46	7.61
		5.32	3.30	0.48	7.54
		5.35	3.31	0.53	7.46
		6.00	3.46	1.26	6.63
		6.12	3.48	1.37	6.57
		6.30	3.50	1.47	6.39
Zinc	3.74	6.40	3.51	1.53	6.30
		5.40	3.46	0.41	7.37
		5.47	3.47	0.44	7.22
		5.43	3.48	0.48	7.16
		5.58	3.49	0.52	7.08
		6.60	3.48	1.42	5.94
		6.73	3.68	1.46	4.83
Manganese	3.81	6.85	3.69	1.52	5.73
		6.00	3.55	0.30	6.32
		6.15	3.57	0.31	6.09
		7.40	3.64	0.33	5.59
		7.80	3.66	0.48	4.23
		7.94	3.67	0.58	4.11
		8.20	3.69	0.78	3.91
		8.55	3.72	1.08	3.70
		8.80	3.74	1.29	3.55
		9.00	3.76	1.50	3.51
		9.10	3.77	1.60	3.50

(c) 8-Hydroxyquinoline 5-Sulphonic Acid

A typical set of experimental titration curves of the chelation with 8-hydroxyquinoline 5-sulphonic acid is given in Figure 4 and the corresponding formation curves in Figure 5, and a summary of the stability constants is given in Table 4.

## V. LITERATURE SURVEY

The only compounds that appear to have been previously investigated quantitatively are those of copper with glycine and alanine, and cobalt with histidine.

Several investigators (11-14) have shown that one, two, or three molecules of glycinate or alaninate ions combine with copper depending on whether the concentrations of chelate ion are equimolar, twice molar, or in large excess respectively.

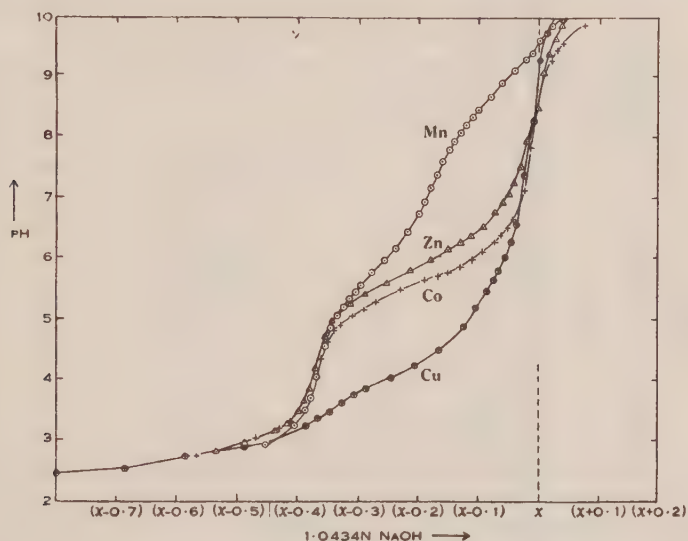


Fig. 2. — The chelation titration curves of metal-histidine complexes. The line (x) represents ml. of NaOH to neutralize the excess acid plus the moles of hydrogen ion liberated during the chelation.

T)

With Cu  $x = 3.685$  ml. of 1.0434N NaOH

Co  $x = 3.57$  " " " "

Zn  $x = 3.742$  " " " "

Mn  $x = 3.808$  " " " "

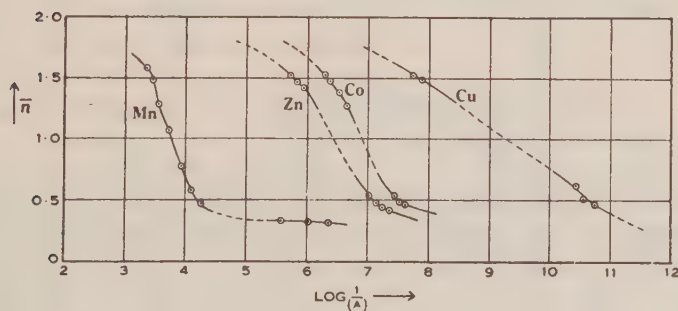


Fig. 3.—Formation curves of metal-histidine complexes.

Polarographic measurements(14) of the overall constants of the *bis* and *tris*





complexes of glycine and alanine with copper gave the following results :

	Log $K_2$	Log $K_3$
Glycine	15.1 (15.42)	16.3
	15.3	16.9*
Alanine	15.0 (14.83)	—

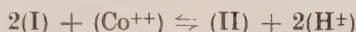
The values obtained in the present work are shown in brackets.

The following results† for the stability of the complex copper glycine have also been reported(15) :

$$\log k_1 = 8.50, \log k_2 = 6.83, \log k_3 = 0.7.$$

These values of  $k_1$  and  $k_2$  agree closely with those (8.51 and 6.91 respectively) given in Table 1.

The only other measurement reported(2) is that on cobalt histidine giving the equilibrium constant for the reaction



in the form

$$\frac{(\text{II})(\text{H}^+)^2}{(\text{Co}^{++})(\text{I})^2} = 4 \times 10^{-7} \text{ at } 24^\circ\text{C.},$$

where



and



This equilibrium constant cannot, however, be directly compared with the value  $K_2$  given in Table 3 until it has been converted to an equivalent form.

On expressing the above equilibrium in terms of the symbols used in this paper it becomes

$$\frac{(\text{CoA}_2)(\text{H}^+)^2}{(\text{Co}^{++})(\text{A}-\text{H}^+)^2} = 4 \times 10^{-7}.$$

On converting this into an expression for  $K_2$  it becomes

$$\frac{(\text{CoA}_2)}{(\text{Co}^{++})(\text{A}^-)^2} = \frac{4 \times 10^{-7}}{(k_{\text{NH}_2})^2} = K_2.$$

From which the calculated value of  $\log K_2$  is 11.96. This value is much smaller than the figure (13.86) given in Table 3.

## VI. DISCUSSION

Throughout Tables 2, 3, and 4 it will be noted that the order of the stability of the metal complexes formed from the different chelating molecules studied is substantially the same as that reported previously(6, 17, 18) and that with the simple amino acids there is also, for any one metal, a small decrease in the stability as the  $\alpha$ -chain is increased in length.

\*Calculated (14) from concentration cell data (13).

†The values quoted here are those reported in *Chemical Abstracts* (15) since the original paper (16) of Flood and Loras was not available.

The chelation of histidine is of interest because it has been found (2) that some metals, especially cobalt, are effective inhibitors of the growth and respiration of various aerobic and anaerobic microorganisms in animal tissues and tumors and that the physiological action of the metal can be overcome by histidine alone among the amino acids tested(3, 4). It is suggested that this effect of histidine may be at least partly concerned with the fact that, because it forms more stable complexes with cobalt than any of the  $\alpha$ -amino acids, it can more readily sequester the cobalt in a form in which it can no longer exercise its inhibitory action.

TABLE 3  
THE STABILITY CONSTANTS OF HISTIDINE COMPLEXES IN WATER AT 25°C.  
Metal Concentration = 0.001M. Chelate Concentration = 0.002M

Metal	Log $k_1$	Log $k_2$	Log $K_2$
Cu	10.56	7.77	18.33
Co	7.52	6.34	13.86
Zn	7.14	5.74	12.88
Mn	4.23	3.51	7.74

TABLE 4  
THE STABILITY CONSTANTS OF 8-HYDROXYQUINOLINE 5-SULPHONIC ACID COMPLEXES IN WATER AT 25°C.  
Metal Concentration = 0.001M. Chelate Concentration = 0.002M

Metal	Log $k_1$	Log $k_2$	Log $K_2$
Cu	11.5	10.7	22.2
Co	9.25	7.45	16.70
Zn	8.70	7.20	15.90

The main interest of the measurements on the complexes of 8-hydroxyquinoline 5-sulphonic acid is that although these complexes are so very much more soluble in water than those of 8-hydroxyquinoline itself, both series of complexes are roughly of the same order of stability.

The values of the stability constants for complexes of 8-hydroxyquinoline reported in Part I were obtained using a 70 per cent. dioxan-water mixture as a solvent. In this earlier work it was also observed that as the percentage of dioxan in the dioxan-water mixture was decreased, there was a decrease in the stability constant.

To render the two series of measurements comparable allowance must be made for the effect of the solvent.

When this is done\*, the values for corresponding compounds are of the same

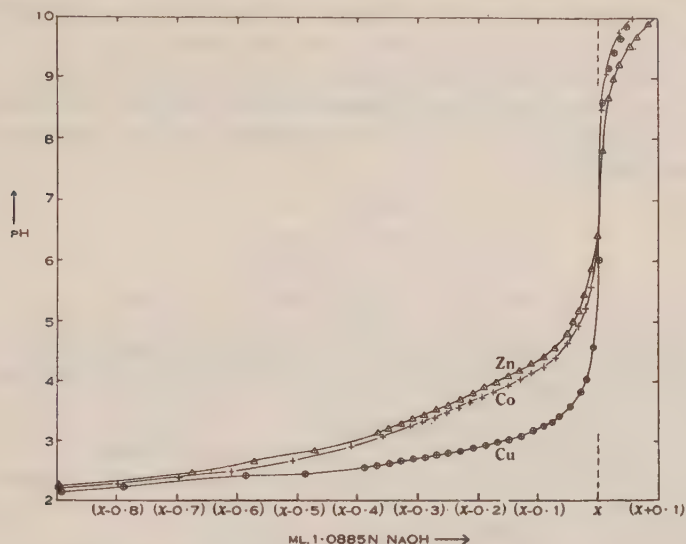


Fig. 4.—The chelation titration curves of metal-8-hydroxyquinoline 5-sulphonic acid complexes. The line (x) represents ml. of NaOH to neutralize the excess acid plus the moles of hydrogen ions liberated during the chelation.

With Cu  $x = 4.294$  ml. of 1.0885N NaOH

Co  $x = 3.515$  " " " "

Zn  $x = 3.177$  " " " "

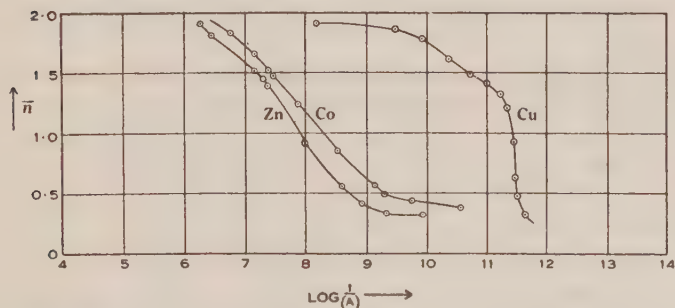


Fig. 5.—Formation curves of metal-8-hydroxyquinoline 5-sulphonic acid complexes.

\*The  $pK_a$  values of 8-hydroxyquinoline are

					$pK_{OH}$	$pK_{NH}$
In water	...	...	...	...	9.45	5.03
In 70% dioxan-water	...	...	...	...	12.33	3.18

The dissociation of 8-hydroxyquinoline in 70 per cent. dioxan-water is notably smaller than in water. A corresponding decrease in the dissociation of the metal complex would be expected. From the variation in the  $pK$  values noted above, the calculated stability constants of 8-hydroxyquinoline complexes in water would be at least 3.4  $pK$  units less than the values in 70 per cent. dioxan-water.

order of magnitude. On the other hand there is a very marked difference in the bacteriostatic properties of the two chelating molecules(12). It may well be that the difference in water solubility of the two series of metal complexes (derived from 8-hydroxyquinoline and 8-hydroxyquinoline 5-sulphonic acid respectively) is a dominant factor in determining bacteriostatic action, and that the greater effectiveness of 8-hydroxyquinoline is due to its precipitated metal complexes remaining on the bacterial cell wall.

## VII. REFERENCES

- (1) VICKERY, H. B., and SCHMIDT, C. L. A.—*Chem. Rev.* **9** : 169 (1931).
- (2) BURK, D., HEARSON, J., CAROLINE, L., and SCHADE, A.—*J. Biol. Chem.* **165** : 723 (1946).
- (3) BURK, D., SCHADE, A., HESSELBACK, M., and FISHER, C.—*Federation Proc.* **5** : 126 (1946).
- (4) BURK, D., HESSELBACK, M., FISHER, C., HEARSON, J., and SCHADE, A.—*Cancer Res. (Abstract)* **6** : 497 (1946).
- (5) HARRIS, L.—*J. Biochem.* **24** : 1080 (1930).
- (6) MALEY, L. E., and MELLOR, D. P.—*Aust. J. Sci. Res. A* **2** : 92 (1949).
- (7) SCHMIDT, C. L. A.—“Chemistry of Amino Acids and Proteins,” p. 613. (Charles C. Thomas : Springfield, 1945.)
- (8) BRUNINGS, K. D.—*J. Amer. Chem. Soc.* **69** : 201 (1947).
- (9) “Handbook of Chemistry and Physics,” 28th Ed. (Chemical Rubber Publishing Co. : Cleveland, 1944.)
- (10) ALBERT, A., and MAGRATH, D.—*Biochem. J.* **41** : 541 (1947).
- (11) BOORSOOK, H., and THIMAN, K. V.—*J. Biol. Chem.* **98** : 671 (1932).
- (12) GOULD, R. K., and VOSBURG, W.—*J. Amer. Chem. Soc.* **64** : 1630 (1942).
- (13) RILEY, H. L., and GALLAFERT, V.—*J. Chem. Soc.* **1931** : 2029 (1931).
- (14) KEFFER, R.—*J. Amer. Chem. Soc.* **68** : 2329 (1946).
- (15) FLOOD, H. L., and LORAS, V.—*Chem. Abstr.* **41** : 6488 (1947).
- (16) FLOOD, H. L., and LORAS, V.—*Tids. Kjemi, Bergvesen Met.* **4** : 35-40 (1944); *Chem. Zentr.* **2** : 987 (1944).
- (17) MELLOR, D. P., and MALEY, L.—*Nature* **159** : 370 (1947).
- (18) MELLOR, D. P., and MALEY, L.—*Ibid.* **161** : 436 (1948).



# THE IODINATION OF AROMATIC COMPOUNDS

## III. HALOGENATION OF AROMATIC ETHERS IN THE PRESENCE OF HYDROGEN PEROXIDE

By L. JURD\*

[Manuscript received July 15, 1949]

### Summary

Aromatic ethers are rapidly iodinated when treated in alcoholic solution with iodine and hydrogen peroxide in the presence of a strong mineral acid. High yields of the mono-iodo-derivatives are obtained with anisole and the 2-naphthyl ethers whilst both mono- and diiodination occur readily with the dimethoxybenzenes. In the absence of a mineral acid the reaction is very slow and only poor yields of the iodo-ethers result.

A convenient process is described for the preparation of *p*-chloro- and *p*-bromo-anisole, and other chloro- and bromo-ethers, the ether being treated with hydrochloric acid and with potassium bromide and sulphuric acid respectively in the presence of hydrogen peroxide.

### I. INTRODUCTION

In a note in 1927 Marsh(1) described the preparation of 2,4,6-triiodophenol by the addition of hydrogen peroxide to an alcoholic solution of iodine and phenol. The reaction has not been investigated further except in 1938 when Kumler(2) used Marsh's process, without modification, to prepare ethyl iodoacetoacetate and other iodo-enols.

It is now found that aromatic ethers cannot be iodinated satisfactorily by Marsh's process, the phenolic ethers being recovered almost quantitatively whilst only small yields of highly-coloured iodo-compounds are obtained from the 2-naphthyl ethers. The use of potassium iodide and sulphuric acid increases the yields, whilst addition of concentrated sulphuric, nitric, or hydrochloric acids to the iodine, hydrogen peroxide, and aromatic ether results in high yields of easily purified iodo-ethers.

The rate of the iodination is accelerated very considerably under these strongly acid conditions, the reaction being violent with most of the ethers investigated. Although anisole and the 2-naphthyl ethers are not diiodinated, *ortho*- and *para*-dimethoxybenzene readily yield both the mono-iodo- and diiodo compounds. This procedure is to be recommended as a preparative method since the usual process of iodination employing iodine and mercuric oxide often gives products contaminated with mercuric iodide(3).

Anisole and the 2-naphthyl ethers are readily monobrominated by means of potassium bromide, sulphuric acid, and hydrogen peroxide, whilst the dimethoxy benzenes form the monobromo- and dibromo-derivatives with one and two molecular equivalents of potassium bromide respectively. As a preparative procedure for

\*Massey Agricultural College, University of New Zealand, Palmerston North, New Zealand.

TABLE I

IODINATION OF ETHERS IN THE PRESENCE OF HYDROGEN PEROXIDE

Ether	Reagents	Reaction Period (min.)	Product	Yield (%)	Iodine		Melting Point	
					Calculated (%)	Found (%)	(°C.)	(lit. °C.)
Anisole	$I_2, H_2SO_4$	20	4-Iodoanisole	80	54.3	54.5	50	50-51
Anisole	$I_2, HCl$	25	4-Iodoanisole	77			50-51	
Anisole	$I_2, HNO_3$	30	4-Iodoanisole	75			50	
	$KI, H_2SO_4$	40	4-Iodoanisole	61			51	
	$I_2$	30	4-Iodoanisole	8			43	
			(Anisole)	(81)				
2-Methoxynaphthalene	$I_2, H_2SO_4$	10	1-Iodo-2-methoxynaphthalene	92	44.7	44.4	88	88
2-Methoxynaphthalene	$I_2, HNO_3$	10	1-Iodo-2-methoxynaphthalene	90			88	
2-Methoxynaphthalene	$I_2, HCl$	10	1-Iodo-2-methoxynaphthalene	90			88	
2-Methoxynaphthalene	$I_2, KCl, H_2SO_4$	30	1-Iodo-2-methoxynaphthalene	93			88	
2-Methoxynaphthalene	$I_2, CH_3COOH$	30	1-Iodo-2-methoxynaphthalene	73			88	
	$KI, H_2SO_4$	30	1-Iodo-2-methoxynaphthalene	80			88	
	$I_2$	30	1-Iodo-2-methoxynaphthalene	70			88	
2-Ethoxynaphthalene	$I_2, H_2SO_4$	10	1-Iodo-2-ethoxynaphthalene	83	42.6	42.6	75	75
2-Ethoxynaphthalene	$KI, H_2SO_4$	30	1-Iodo-2-ethoxynaphthalene	77			74	
2-Ethoxynaphthalene	$I_2$	30	1-Iodo-2-ethoxynaphthalene	67			74	
2-Benzoxynaphthalene	$I_2, H_2SO_4$	40	1-Iodo-2-benzoxynaphthalene	61	35.3	34.6	86	86
Veratrole	$I_2, H_2SO_4$	40	4-Iodo-2-benzoxynaphthalene	57	48.1	48.5	174-175*	
1,4-Dimethoxybenzene	$I_2, H_2SO_4$	45	4-Iodo-2-benzoxynaphthalene (1,4-Dimethoxybenzene)	(34)				
			2-Iodo-1,4-dimethoxybenzene	45	48.1	47.3	180-183**	
4-Iodoveratrole	$I_2, H_2SO_4$	30	4,5-Diiodoveratrole	71	65.1	65.0	131-135	132
2-Iodo-1,4-dimethoxybenzene	$I_2, H_2SO_4$	25	2,5-Diiodo-1,4-dimethoxybenzene	61	65.1	64.8	172-173	171

\* b.p./30 mm. \*\* b.p./40 mm.

TABLE 2  
BROMINATION OF ETHERS IN THE PRESENCE OF HYDROGEN PEROXIDE

Ether	Reagents	Reaction Period (min.)	Products	Yield (%)	Bromine		Melting Point	
					Calculated (%)	Found (%)	(°C.)	(lit. °C.)
Anisole	KBr, H <sub>2</sub> SO <sub>4</sub>	25	4-Bromoanisole	80	42.7	42.7	215-217*	215
1,4-Dimethoxybenzene	I <sub>2</sub> , KBr, H <sub>2</sub> SO <sub>4</sub>	20	4-Bromoanisole	70			214-220	
	KBr(1M), H <sub>2</sub> SO <sub>4</sub>	30	2-Bromo-1,4-dimethoxybenzene	66	36.8	36.6	164-166**	
	KBr(2M), H <sub>2</sub> SO <sub>4</sub>	20	2,5-Dibromo-1,4-dimethoxybenzene	76	54.0	54.3	143	142
	I <sub>2</sub> , KBr(2M), H <sub>2</sub> SO <sub>4</sub>	30	2,5-Dibromo-1,4-dimethoxybenzene	45			137	
	I <sub>2</sub> , HBr	30	2,5-Dibromo-1,4-dimethoxybenzene	50	54.6	53.6	137	
2-Methoxynaphthalene	KBr, H <sub>2</sub> SO <sub>4</sub>	30	1-Bromo-2-methoxynaphthalene	80	33.7	33.4	84-85	84-85
2-Benzoxynaphthalene	I <sub>2</sub> , KBr, H <sub>2</sub> SO <sub>4</sub>	30		65			83-85	
	KBr, H <sub>2</sub> SO <sub>4</sub>	30		78	25.5	24.9	106	
4-Bromoanisole	KBr, H <sub>2</sub> SO <sub>4</sub>	20	1(?) -Bromo-2-benzoxynaphthalene ***	(72)				
	HBr	20	(4-Bromoanisole) (4-Bromoanisole)	(78)				

\* b.p./atm. pres. \*\* b.p./55 mm. \*\*\* compound not previously described.

TABLE 3  
CHLORINATION OF ETHERS IN THE PRESENCE OF HYDROGEN PEROXIDE

Ether	Reagents	Reaction Period (min.)	Products	Yield (%)	Melting Point	
					(°C.)	(lit. °C.)
Anisole	KCl, H <sub>2</sub> SO <sub>4</sub>	20	(Anisole)	(86)		
Anisole	HCl	50	4-Chloroanisole	79	200-205*	
2-Methoxynaphthalene	KCl, H <sub>2</sub> SO <sub>4</sub>	30	1-Chloro-2-methoxynaphthalene	80	67	67-69
4-Iodoanisole	KCl, H <sub>2</sub> SO <sub>4</sub>	30	(4-Iodoanisole)	(87)		
1-Iodo-2-methoxynaphthalene	KCl, H <sub>2</sub> SO <sub>4</sub>	30	(1-Iodo-2-methoxynaphthalene)	(91)		

\* b.p./atm. pres.

*p*-bromoanisole it is to be preferred to that employing elementary bromine which produces with an equivalent quantity of anisole a mixture of 4-bromo- and 2, 4-dibromoanisole.

Although *p*-dimethoxybenzene and 2-methoxynaphthalene are readily chlorinated by means of potassium chloride, sulphuric acid, and hydrogen peroxide, anisole does not react appreciably under these conditions. However, 80 per cent. yields of *p*-chloroanisole result when anisole is heated with excess of hydrogen chloride in the presence of hydrogen peroxide.

In general, addition of potassium chloride to an iodine-sulphuric acid mixture does not decrease the yields of the iodo-ethers, whereas in the presence of potassium bromide, brominated ethers are chiefly produced. A parallel might be drawn between these two processes and the processes of iodination and bromination with the reagents iodine monochloride and iodine monobromide respectively, described by Miltzer(4). This aspect of the work is at present under investigation.

The results obtained in the halogenation of the ethers by various reagents are summarized in Tables 1-3.

## II. EXPERIMENTAL

The following procedures illustrate the nature of the processes summarized in Tables 1-3.

### (a) 4-Iodoanisole

Hydrogen peroxide (10.0 cc., 30%) was added in small portions during 20 minutes to a mixture of anisole (5.93 g.), powdered iodine (6.97 g.), and concentrated sulphuric acid (1.0 cc., sp. gr. 1.83) in alcohol (30 cc., 95%), the reaction mixture being warmed gently and vigorously agitated between each addition. After adding the hydrogen peroxide solution (c.5 cc.) a violent reaction occurred, this being allowed to subside before completing the addition. The 4-iodoanisole, which separated as a dark oil on adding excess water to the cooled reaction mixture, was extracted with ether. After washing with dilute aqueous sodium thiosulphate and water, and drying over anhydrous calcium chloride, the solvent was distilled from the ether extract. On standing in the refrigerator, the residual red oil rapidly solidified to an orange, crystalline mass, m.p. 49°C. (11.2 g.). Recrystallization of this product from 90% ethanol gave light yellow plates, m.p. 50°C. (10.18 g., 80%).

### (b) 4-Bromoanisole

A vigorously agitated mixture of anisole (7.78 g.), concentrated sulphuric acid (5.0 cc., sp. gr. 1.83), powdered potassium bromide (8.57 g.), and alcohol (30 cc.) was warmed and treated with hydrogen peroxide solution (12.0 cc., 30%) added gradually during 20 minutes. Water (10.0 cc.) was then added, the remaining potassium bromide immediately went into solution and a violent reaction occurred. After standing for 5 minutes to ensure completion of the reaction, excess water was added to precipitate the product. After cooling the yellow oil which separated was extracted with ether, the ethereal solution being washed with water and dried with anhydrous calcium chloride. The solvent was then removed and the residual oil distilled at atmospheric pressure, the colourless, 4-bromoanisole fraction, b.p. 215-217°C. (lit. 215°C.), being collected (10.79 g., 80%).

### (c) 4-Chloroanisole

A solution of anisole (7.4 g.) and concentrated hydrochloric acid (15.0 cc., 31.8%) in alcohol (40.0 cc.) was refluxed gently on the water bath and treated with hydrogen peroxide (20.0 cc., 30%) during 30 minutes. Refluxing was continued for a further 20 minutes, the mixture then being cooled and treated with excess water. The oil which precipitated was separated, dried over calcium chloride, and distilled at atmospheric pressure, the 4-chloroanisole fraction, b.p. 200-205°C. (lit. 200°C.), being collected (7.7 g., 79%).



(d) 4,5-Diiodo-1, 2-Dimethoxybenzene

A flask containing a mixture of 4-iodoveratrole (4.12 g.), powdered iodine (2.0 g.) concentrated sulphuric acid (1.0 cc., sp. gr. 1.83), and alcohol (10.0 cc.) was fitted with a reflux condenser and suspended over a hot water-bath so that the temperature of the reaction mixture was maintained above 70°C. Hydrogen peroxide (5.0 cc., 30%) was added through the condenser during 15 minutes, the violent reaction which occurred with each addition being allowed to subside before making the next addition. The mixture was refluxed gently for 15 minutes, during which time the colour of the iodine was completely discharged. On adding water (50.0 cc.) an oil separated which rapidly solidified to a crystalline mass (5.50 g.). This was collected, washed with water, and recrystallized from alcohol, colourless needles of 4,5-diiodoveratrole, m.p. 131.5°C., being obtained (4.31 g., 71%).

(e) 2,5-Dibromo-1, 4-Dimethoxybenzene

Concentrated sulphuric acid (2.5 cc., sp. gr. 1.83) was added to a mixture of powdered potassium bromide (3.55 g.), *p*-dimethoxybenzene (2.0 g.), water (10.0 cc.), and alcohol (20.0 cc.). The mixture was refluxed gently and treated with hydrogen peroxide (6.0 cc., 30%) during 15 minutes. Refluxing was continued for a further 5 minutes during which time a heavy precipitate of almost colourless needles separated. Excess water was then added gradually to the hot mixture, the crystalline precipitate collected, washed with water, and dried (3.80 g., m.p. 140°C.). Recrystallization of this product from alcohol gave 2, 5-dibromohydroquinone dimethyl ether as colourless needles, m.p. 143°C. (3.28 g., 76%).

(f) 1-Iodo-2-Methoxynaphthalene

Hydrogen peroxide (5.0 cc., 30%) was added in 1.0 cc. portions during 5 minutes to a warm, agitated mixture of 2-methoxynaphthalene (2.0 g.), concentrated sulphuric acid (1.0 cc., sp. gr. 1.83), and powdered iodine (1.64 g.) in alcohol (15.0 cc.). The violent reaction which occurred after two or three minutes was allowed to subside before completing the addition of peroxide. Agitation and warming were continued for a further 5 minutes during which time a yellow, crystalline precipitate of the iodo-compound separated from the orange solution. Water (30 cc.) was then added gradually and with shaking to the hot mixture. After cooling in the refrigerator, the crystalline precipitate (3.50 g.) melting at 87°C., was collected, washed with dilute sodium thiosulphate solution and water, dried, and recrystallized from petroleum ether, almost colourless plates, m.p. 88°C., being obtained (3.27 g., 92%).

(g) 1-Chloro-2-Methoxynaphthalene

A mixture of powdered potassium chloride (0.94 g.), concentrated sulphuric acid (1.0 cc., sp. gr. 1.83), and 2-methoxynaphthalene (2.0 g.) in alcohol (15.0 cc.) was warmed to 60-70°C. Hydrogen peroxide solution (5.0 cc., 30%) was then added in several small portions, with vigorous agitation between each addition, during 15 minutes. Agitation and warming were continued for a further 15 minutes, the chloro-compound separating as a light yellow oil. After adding water (30 cc.) and cooling in the refrigerator, the precipitated 1-chloro-2-methoxynaphthalene solidified to a mass of lustrous, almost colourless plates which were collected, washed with water, and dried (2.12 g., m.p. 62°C.). Plates, melting at 67°C. were obtained on one recrystallization from aqueous alcohol (1.96 g., 80%).

### III. REFERENCES

- (1) MARSH, J. E.—*J. Chem. Soc.* **1927**: 3164 (1927).
- (2) KUMLER, W. D.—*J. Amer. Chem. Soc.* **60**: 855 (1938).
- (3) JURD, L.—*Aust. J. Sci. Res. A* **2**: 246 (1949).
- (4) MILITZER, W.—*J. Amer. Chem. Soc.* **60**: 256 (1938).

# STUDIES ON THE LIGNIN OF *EUCALYPTUS REGNANS*

## IV. ALKALI LIGNIN

By J. W. T. MEREWETHER\*

[Manuscript received July 19, 1949]

### Summary

By pulping *Eucalyptus regnans* using the soda process and acidifying the black liquor a crude alkali lignin has been isolated. This has been separated into two fractions, dioxan-ether insoluble alkali lignin-A and dioxan-ether soluble alkali lignin-B.

A series of derivatives has been made from each. The data for alkali lignin-A are consistent with the empirical formula  $C_{68}H_{78}O_{27}$ , containing nine methoxyl groups, three acidic, and four alcoholic hydroxyl groups, one of which is tertiary, one carbonyl, and one active methylene group.

Alkali lignin-B has been found to be identical with the alkali lignin previously isolated from sulphate black liquor.

## I. INTRODUCTION

Alkali lignin, that is, lignin isolated with alkali and precipitated by means of mineral acids, has been used as a starting material by many investigators, and whatever changes native lignin undergoes during conversion to alkali lignin, the available data show that much of our knowledge of lignin chemistry may be applied to alkali lignin. For example, it shows similar ultraviolet absorption spectra(1, 2), oxidation(3) gives vanillin (and syringaldehyde, with hardwoods), hydrogenation(4) yields derivatives of 4-*n*-propylcyclohexanol, and the isolation of 4-*n*-propylguaiacol(5) shows that this characteristic lignin unit is still present. On the other hand, widely differing data on its elementary composition and the presence of constituent groups have been reported. For example, in Table 1 are shown the results obtained by various authors for different alkali lignins.

Different figures can be expected from lignins of different woods, but in view of the fact that the methods of delignification were substantially the same it is surprising that the spruce alkali lignin of Dorée and Barton-Wright(6) differed so greatly from those of Marshall, Brauns, and Hibbert(7).

## II. *EUCALYPTUS REGNANS* ALKALI LIGNIN

The object of the work described in the present paper was to obtain data on the chemical nature of *E. regnans* alkali lignin. It has been shown that when this wood is pulped by the sulphate process and the spent liquors acidified, the product can be separated into a thioglignin(11, 12) and an alkali lignin(13). These two products presumably result from the interaction of lignin with sodium hydrosulphide and hydroxide respectively, and it should be possible to produce the same alkali

\*Research Laboratories, Australian Paper Manufacturers Limited, Melbourne.

lignin if the same wood is treated with sodium hydroxide alone. Secondly, a comparison of the chemical properties of the lignin derivatives produced during soda and sulphate pulping should yield data of value in interpreting the role of the sodium sulphide in the later process.

TABLE 1

Raw Material	Empirical Formula	Methoxyl Groups : Hydroxyl Groups	Analysis			Ref.
			OCH <sub>3</sub> (%)	C (%)	H (%)	
Spruce ... ..	C <sub>20</sub> H <sub>20</sub> O <sub>6</sub>	2 : 1	18.0	66.9	5.6	(6)
Spruce ... ..	A C <sub>67</sub> H <sub>71</sub> O <sub>23</sub>	6 : 7	15.0	64.8	5.7	(7)
	B C <sub>108</sub> H <sub>107</sub> O <sub>37</sub>	9 : 11	14.0	65.0	5.4	
Spruce ... ..	C <sub>45</sub> H <sub>48</sub> O <sub>16</sub>	4 : 5	—	64.0	5.5	(8)
Rye straw ... ..	C <sub>40</sub> H <sub>44</sub> O <sub>15</sub>	4 : 4	14.3	63.0	5.6	(9)
Wheat straw ...	A C <sub>42</sub> H <sub>48</sub> O <sub>16</sub>	4 : 5	16.6	62.2	5.9	(10)
	B C <sub>40</sub> H <sub>42</sub> O <sub>16</sub>	4 : 4	17.3	61.8	5.4	

TABLE 2  
ALKALI LIGNIN-B

Compound		C (%)	H (%)	OCH <sub>3</sub> (%)	COCH <sub>3</sub> (%)
Alkali lignin-B ... ..	f.	62.7	6.0	25.1	—
	c.	63.0	6.0	24.7	—
Hexamethyl alkali lignin-B ... ..	f.	64.2	6.3	33.7	—
	c.	64.1	6.4	33.8	—
Octoacetyl alkali lignin-B ... ..	f.	62.3	5.7	21.6	16.8
	c.	62.1	5.8	21.1	16.5
Triacetylpentamethyl alkali lignin-B ... ..	f.	63.2	6.5	30.1	6.7
	c.	63.4	6.2	30.2	6.6
Trityl alkali lignin-B ... ..	f.	67.0	6.0	21.8	—
	c.	66.8	6.0	21.8	—

A sample of the same wood was therefore pulped by the soda process, using conditions recommended by A. Meller (personal communication) and the crude alkali lignin isolated by acidifying the spent liquor. On treatment with anhydrous

dioxan this yielded a dioxan-insoluble fraction which, when distilled with 12 per cent. hydrochloric acid, yielded furfural equivalent to 70 per cent. xylan. This was presumably hemicellulose dissolved by the drastic action of the unbuffered sodium hydroxide. From the dioxan solution one alkali lignin fraction was obtained by precipitation into anhydrous ether, and a second fraction resulted by adding the dioxan-ether mother liquors to an equal volume of light petroleum. Repeated purification finally yielded these two fractions in a pure state, and they were designated alkali lignin-A and alkali lignin-B respectively.

TABLE 3  
ALKALI LIGNIN-A AND THIOLIGNIN

Compound	C (%)	H (%)	N (%)	S (%)	OCH <sub>3</sub> (%)	Acyl (%)
Alkali lignin-A ... ..	61.5	6.0	—	—	20.9	—
Thiolignin ... ..	61.5	5.7	—	2.3	21.1	—
Trimethyl alkali lignin-A ...	62.7	6.2	—	—	27.7	—
Trimethyl thiolignin ... ..	62.4	6.1	—	2.2	27.1	—
Heptacetyl alkali lignin-A ...	61.0	5.7	—	—	17.6	18.3
Heptacetyl thiolignin ... ..	61.0	5.6	—	2.2	18.4	18.6
Tetracetyltrimethyl alkali lignin-A	61.8	6.2	—	—	24.5	11.2
Tetracetyltrimethyl thiolignin ...	62.1	5.9	—	2.0	24.9	11.0
Heptabenzoyl alkali lignin-A ...	68.4	5.2	—	—	13.3	35.6
Heptabenzoyl thiolignin ... ..	68.5	5.0	—	1.7	13.5	35.9
Hexabenzoyl alkali lignin-A ...	67.6	5.4	—	—	14.4	32.1
Hexabenzoyl thiolignin ... ..	67.6	5.3	—	1.5	14.0	31.8
Benzal alkali lignin-A ... ..	64.1	5.9	—	—	19.4	—
Benzal thiolignin ... ..	63.4	5.6	—	2.4	19.9	—
<i>m</i> -Nitrobenzal alkali lignin-A ...	62.0	5.5	1.0	—	19.1	—
<i>m</i> -Nitrobenzal thiolignin ... ..	61.8	5.3	1.1	2.1	19.4	—
Alkali lignin-A phenylosazone ...	64.0	5.7	3.7	—	18.6	—
Thiolignin phenylosazone ... ..	64.3	5.5	3.6	2.1	18.7	—
Alkali lignin-A <i>p</i> -nitrophenylhydrazone	60.8	5.7	3.1	—	19.2	—
Thiolignin <i>p</i> -nitrophenylhydrazone	61.0	5.4	3.0	2.1	19.2	—

As expected, alkali lignin-B appeared to be identical with the dioxan-ether soluble alkali lignin previously isolated from the sulphate pulping of this wood, and this was confirmed by the preparation and analysis of four derivatives (Table 2).

Turning now to alkali lignin-A, it is interesting to note that it is strikingly similar to thiolignin. Chemical analysis gave carbon, hydrogen, and methoxyl



figures very close to those of thioglignin, and potentiometric titrations gave a curve of the same general type as obtained with thioglignin; there was a pronounced point of inflection at pH 7.6, corresponding to an equivalent weight of 670, a value of the same order as that obtained for thioglignin. Nine derivatives were made, and the analyses of these (Table 3) show a marked similarity to the corresponding thioglignin compounds, indicating that the two compounds must have a very similar structure.

TABLE 4  
ALKALI LIGNIN-A AND DERIVATIVES

Compound		C (%)	H (%)	N (%)	OCH <sub>3</sub> (%)	Acyl (%)
Alkali lignin-A ... ..	f.	61.5	6.0	—	20.9	—
	c.	61.6	5.9	—	21.1	—
Trimethyl alkali lignin-A ...	f.	62.7	6.2	—	27.7	—
	c.	62.3	6.2	—	27.2	—
Heptacetyl alkali lignin-A ...	f.	61.0	5.7	—	17.6	18.3
	c.	60.7	5.7	—	17.2	18.5
Tetracetyltrimethyl alkali lignin-A	f.	61.8	6.2	—	24.5	11.2
	c.	61.7	6.0	—	24.2	11.2
Heptabenzoyl alkali lignin-A ...	f.	68.4	5.2	—	13.3	35.6
	c.	68.4	5.2	—	13.6	35.8
Hexabenzoyl alkali lignin-A ...	f.	67.6	5.4	—	14.4	32.1
	c.	67.7	5.3	—	14.3	32.2
Benzal alkali lignin-A ... ..	f.	64.1	5.9	—	19.4	—
	c.	63.7	5.8	—	19.7	—
<i>m</i> -Nitrobenzal alkali lignin-A ...	f.	62.0	5.5	1.0	19.1	—
	c.	61.8	5.6	1.0	19.1	—
Alkali lignin-A phenylosazone ...	f.	64.0	5.7	3.7	18.6	—
	c.	63.9	5.9	3.7	18.6	—
Alkali lignin-A <i>p</i> -nitrophenylhydrazone	f.	60.8	5.7	3.1	19.2	—
	c.	60.7	5.7	2.9	19.1	—

The solubilities of alkali lignin-A and its derivatives in various solvents are likewise practically identical with those of the corresponding thioglignin compounds, but with one important exception. Thioglignin is precipitated at pH 3.7, alkali lignin-A at pH 4.1, on acidification of their respective alkaline solutions, suggesting that the greater acidity of thioglignin is due to the presence of the mercaptan group.

The formula  $C_{68}H_{76}O_{25}S$  was found to be consistent with all the data for thiolignin. In view of the similarity in chemical composition alkali lignin-A must have a similar structure. A rough approximation for the minimum molecular weight may be calculated from the nitrogen content of the condensation product with *m*-nitrobenzaldehyde (1.0 per cent.) and this value (1400) is in reasonable agreement with that obtained by doubling the equivalent weight (1340). The empirical formula  $C_{68}H_{78}O_{27}(1327)$ , or  $C_{59}H_{51}O_{18}(OCH_3)_9$ , has been tentatively selected as the formula that agrees best with the data, and a comparison of the analytical data with those calculated for this formula is given in Table 4.

Like thiolignin, alkali lignin-A has seven hydroxyl groups which can be acetylated with acetic anhydride in pyridine, and three of these are sufficiently acidic to react with diazomethane. The four alcoholic hydroxyls react with benzoyl chloride in pyridine, but only three react in the presence of aqueous alkali, indicating the presence of a tertiary hydroxyl group. The presence of a carbonyl group is indicated by the formation of a *p*-nitrophenylhydrazone, and as phenylhydrazine yields a phenylosazone one of the alcoholic hydroxyls must be attached to the adjacent carbon atom. Finally the presence of an active methylene group is shown by the formation of condensation products with benzaldehydes.

TABLE 5  
SPRUCE AND EUCALYPTUS REGNANS LIGNINS

Spruce alkali lignin-A	$C_{61}H_{46}O_{16}(OCH_3)_9(OH)_7$
<i>E. regnans</i> alkali lignin-A	$C_{59}H_{44}O_{11}(OCH_3)_9(OH)_7$
<i>E. regnans</i> thiolignin	$C_{59}H_{42}O_{16}(OCH_3)_9(OH)_6SH$

It is of interest to note that not only is *E. regnans* alkali lignin-A very similar to thiolignin, but a comparison of the data with those of Marshall, Brauns, and Hibbert(7) shows that there is a close resemblance between the alkali lignins-A of spruce and *E. regnans*. Not only has each molecule seven hydroxyl groups, but in both cases three are acidic and four alcoholic. While the molecule of spruce alkali lignin-A has six methoxyl groups compared with nine for *E. regnans*, this is readily understood when one remembers that hardwood lignin is known to contain syringyl units as well as guaiacyl. In Table 5 are set out the empirical formulae of these three compounds, and it is seen that, on a methoxyl free basis, the molecules are very similar.

## II. EXPERIMENTAL

### (a) Crude Alkali Lignin

Six logs of green 21-year-old *E. regnans* from Powelltown, Victoria, were converted into chips, screened to remove oversize and fines, sorted to remove gum veins etc. and air dried. In a steam jacketed rotary iron digester were mixed chips (4000 g. dry weight), sodium hydroxide (1400 g. active NaOH), and water to make the total volume of liquor (including the moisture in the chips) 16 litres. The charge was taken up to the boiling point in 45 minutes when all the air was blown out, to 166 C. in another 45 minutes, and maintained at that temperature for a further 2½ hours. On cooling, the spent liquor was filtered on a Buchner, and the residual pulp washed three times with water.

The combined filtrate and washings were filtered once more, warmed to 50-55 C. and 50%  $H_2SO_4$  added slowly with stirring until the mixture was acid to Congo red. Vigorous foaming

occurred and the crude lignin was precipitated. This was filtered off, washed well with water, and dried in a vacuum desiccator. Yield 830 g.

(b) *Alkali Lignin-A and Alkali Lignin-B*

Treatment of this crude lignin (200 g.) with dioxan (2000 cc.) yielded a dioxan insoluble fraction (28 g.) which appeared to be hemicellulose, the furfural yield corresponding to 70% xylan. The filtered dioxan solution was then worked up according to the methods previously described for the separation of thioglignin and alkali lignin (11, 13), yielding a dioxan-ether insoluble fraction, alkali lignin-A (87 g.), and a dioxan-ether soluble fraction, alkali lignin-B (33 g.). From the residues and intermediate fractions of four such batches were obtained a further 27 g. crude alkali lignin-A, 25 g. crude alkali lignin-B, and 30 g. purified alkali lignin-B.

Found for alkali lignin-A: C, 61.5; H, 6.0;  $\text{OCH}_3$ , 20.9%; eq. wt. 670.

Calculated for  $\text{C}_{68}\text{H}_{78}\text{O}_{27}$ (1327): C, 61.6; H, 5.9;  $\text{OCH}_3$ , 21.1%; eq. wt. 664.

Alkali lignin-A is a fine light brown powder which commences to soften at 183°C., but which has no sharp melting point. It is soluble in acetic acid, pyridine, dioxan, and aqueous sodium hydroxide, slightly soluble in ethyl alcohol and acetone, and insoluble in ether, light petroleum, chloroform, ethyl acetate, benzene, water, and aqueous sodium bicarbonate.

Found for alkali lignin-B: C, 62.7; H, 6.0;  $\text{OCH}_3$ , 25.1%.

Calculated for  $\text{C}_{92}\text{H}_{104}\text{O}_{34}$ (1754): C, 63.0; H, 6.0;  $\text{OCH}_3$ , 24.7%.

(c) *Trimethyl Alkali Lignin-A*

Alkali lignin-A (20 g.) was treated with diazomethane under the conditions described for the preparation of trimethylthioglignin (11). After three methylations the product was isolated and purified twice from dioxan-ether. Yield 16.2 g.

Found: C, 62.7; H, 6.2;  $\text{OCH}_3$ , 27.7%.

Calculated for  $\text{C}_{71}\text{H}_{84}\text{O}_{27}$ (1369): C, 62.3; H, 6.2;  $\text{OCH}_3$ , 27.2%.

Trimethyl alkali lignin-A is a yellowish-buff powder which commences to soften at 168°C. It is soluble in chloroform, acetone, acetic acid, pyridine, and dioxan, and insoluble in ether, light petroleum, ethyl alcohol, ethyl acetate, benzene, water, and cold dilute aqueous sodium hydroxide.

(d) *Heptacetyl Alkali Lignin-A*

Alkali lignin-A (10 g.) was acetylated under the conditions described for the preparation of heptacetylthioglignin (11), and purified twice from dioxan-ether. Yield 8 g.

Found: C, 61.0; H, 5.7;  $\text{OCH}_3$ , 17.6;  $\text{COCH}_3$ , 18.3%.

Calculated for  $\text{C}_{82}\text{H}_{92}\text{O}_{34}$ (1622): C, 60.7; H, 5.7;  $\text{OCH}_3$ , 17.2;  $\text{COCH}_3$ , 18.5%.

Heptacetyl alkali lignin-A is a greyish-buff powder which commences to soften at 169°C. It is soluble in chloroform, acetone, acetic acid, ethyl acetate, benzene, pyridine, and dioxan, and insoluble in ether, light petroleum, ethyl alcohol, and cold dilute aqueous sodium hydroxide.

(e) *Tetracetyltrimethyl Alkali Lignin-A*

Trimethyl alkali lignin-A (5 g.) was acetylated under the conditions described for the preparation of tetracetyltrimethylthioglignin (11), and purified twice from dioxan-ether. Yield 4.6 g.

Found: C, 61.8; H, 6.2;  $\text{OCH}_3$ , 24.5;  $\text{COCH}_3$ , 11.2%.

Calculated for  $\text{C}_{79}\text{H}_{92}\text{O}_{31}$ (1538): C, 61.7; H, 6.0;  $\text{OCH}_3$ , 24.2;  $\text{COCH}_3$ , 11.2%.

Tetracetyltrimethyl alkali lignin-A is a dull yellow powder which commences to soften at 157°C., and is soluble in the same solvents as heptacetyl alkali lignin-A.

(f) *Heptabenzoyl Alkali Lignin-A*

Alkali lignin-A (6 g.) was benzoylated under the conditions described for the preparation of heptabenzoylthioglignin (12), and purified once from chloroform-ether-light petroleum and twice from dioxan-ether. Yield 5.4 g.

Found: C, 68.4; H, 5.2;  $\text{OCH}_3$ , 13.3;  $\text{COC}_6\text{H}_5$ , 35.6%.

Calculated for  $\text{C}_{117}\text{H}_{106}\text{O}_{34}$ (2056): C, 68.4; H, 5.2;  $\text{OCH}_3$ , 13.6;  $\text{COC}_6\text{H}_5$ , 35.8%.



Heptabenzoyl alkali lignin-A is a brownish powder which begins to soften at 181°C. It is soluble in the same solvents as heptacetyl alkali lignin-A except that it is only slightly soluble in ethyl acetate.

(g) *Hexabenzoyl Alkali Lignin-A*

Alkali lignin-A (6 g.) was benzoylated under the conditions described for the preparation of hexabenzoylthiolignin (12), and purified once from dioxan-water and twice from dioxan-ether. Yield 4.3 g.

Found: C, 67.6; H, 5.4;  $\text{OCH}_3$ , 14.4;  $\text{COC}_6\text{H}_5$ , 32.1%.

Calculated for  $\text{C}_{110}\text{H}_{102}\text{O}_{33}$  (1952): C, 67.7; H, 5.3;  $\text{OCH}_3$ , 14.3;  $\text{COC}_6\text{H}_5$ , 32.2%.

Hexabenzoyl alkali lignin-A is a brownish powder which begins to soften at 178°C. and is soluble in the same solvents as heptacetyl alkali lignin-A.

(h) *Alkali Lignin-A Phenylsazone*

Alkali lignin-A (5 g.) was treated with phenylhydrazine under the conditions described for the preparation of thiolignin phenylsazone (12), purified twice from dioxan-ether, and finally extracted with ether in a soxhlet extractor. Yield 4.1 g.

Found: C, 64.0; H, 5.7; N, 3.7;  $\text{OCH}_3$ , 18.6%.

Calculated for  $\text{C}_{80}\text{H}_{88}\text{O}_{25}\text{N}_4$  (1506): C, 63.9; H, 5.9; N, 3.7;  $\text{OCH}_3$ , 18.6%.

Alkali lignin-A phenylsazone is a dark brown powder that begins to soften at 200°C. It is soluble in ethyl alcohol, acetone, acetic acid, pyridine, dioxan, and aqueous sodium hydroxide, slightly soluble in chloroform and ethyl acetate, and insoluble in ether, light petroleum, benzene, and water.

(i) *Alkali Lignin-A, p-Nitrophenylhydrazone*

Alkali lignin-A (5 g.) was treated with *p*-nitrophenylhydrazine under the conditions described for the preparation of thiolignin *p*-nitrophenylhydrazone (12), purified twice from dioxan-ether, and finally extracted with ether in a soxhlet extractor. Yield 4.0 g.

Found: C, 60.8; H, 5.7; N, 3.1;  $\text{OCH}_3$ , 19.2%.

Calculated for  $\text{C}_{74}\text{H}_{82}\text{O}_{28}\text{N}_3$  (1462): C, 60.7; H, 5.7; N, 2.9;  $\text{OCH}_3$ , 19.1%.

Alkali lignin-A *p*-nitrophenylhydrazone is a bright brown powder which commences to soften at 203°C. It is soluble in the same solvents as alkali lignin-A phenylsazone except that it is only slightly soluble in ethyl alcohol and acetone.

(j) *Benzal Alkali Lignin-A*

Alkali lignin-A (4 g.) was treated with benzaldehyde under the conditions described for the preparation of benzalthiolignin (12), and purified twice from dioxan-ether. Yield 2.9 g.

Found: C, 64.1; H, 5.9;  $\text{OCH}_3$ , 19.4%.

Calculated for  $\text{C}_{75}\text{H}_{83}\text{O}_{27}$  (1415): C, 63.7; H, 5.8;  $\text{OCH}_3$ , 19.7%.

Benzal alkali lignin-A is a brown powder which commences to soften at 210°C., is soluble in ethyl alcohol, acetic acid, pyridine, dioxan, and aqueous sodium hydroxide, slightly soluble in acetone, and insoluble in chloroform, ether, light petroleum, ethyl acetate, benzene, and water.

(k) *m-Nitrobenzal Alkali Lignin-A*

Alkali lignin-A (4 g.) was treated with *m*-nitrobenzaldehyde as described for *m*-nitrobenzalthiolignin (12), purified twice from dioxan ether, and extracted with ether in a soxhlet extractor. Yield 3.2 g.

Found: C, 62.0; H, 5.5; N, 1.0;  $\text{OCH}_3$ , 19.1%.

Calculated for  $\text{C}_{75}\text{H}_{81}\text{NO}_{29}$  (1460): C, 61.8; H, 5.6; N, 1.0;  $\text{OCH}_3$ , 19.1%.

*m*-Nitrobenzal alkali lignin-A is a brown powder which begins to soften at 212°C. and is soluble in the same solvents as benzal alkali lignin-A.

(l) *Hexamethyl Alkali Lignin-B*

Alkali lignin-B (14.1 g.) was treated with diazomethane under the conditions previously



described for the preparation of hexamethyl alkali lignin (13), remethylated twice, and purified three times from chloroform-ether-light petroleum. Yield 7.1 g.

Found: C, 64.2; H, 6.3;  $\text{OCH}_3$ , 33.7%.

Calculated for  $\text{C}_{98}\text{H}_{116}\text{O}_{34}$  (1838): C, 64.1; H, 6.4;  $\text{OCH}_3$ , 33.8%.

(m) *Octoacetyl Alkali Lignin-B*

Alkali lignin-B (5 g.) was acetylated under the conditions described for the preparation of octoacetyl alkali lignin (13) and purified three times from chloroform-ether-light petroleum. Yield 2.0 g.

Found: C, 62.3; H, 5.7;  $\text{OCH}_3$ , 21.6;  $\text{COCH}_3$ , 16.8%.

Calculated for  $\text{C}_{108}\text{H}_{130}\text{O}_{42}$  (2090): C, 62.1; H, 5.8;  $\text{OCH}_3$ , 21.1;  $\text{COCH}_3$ , 16.5%.

(n) *Triacetylpentamethyl Alkali Lignin-B*

Hexamethyl alkali lignin-B (2 g.) was acetylated under the conditions described for the preparation of triacetylpentamethyl alkali lignin (13) and purified twice from chloroform-ether-light petroleum. Yield 0.8 g.

Found: C, 63.2; H, 6.5;  $\text{OCH}_3$ , 30.1;  $\text{COCH}_3$ , 6.7%.

Calculated for  $\text{C}_{103}\text{H}_{120}\text{O}_{37}$  (1950): C, 63.4; H, 6.2;  $\text{OCH}_3$ , 30.2;  $\text{COCH}_3$ , 6.6%.

(o) *Trityl Alkali Lignin-B*

Alkali lignin-B (3 g.) was treated with triphenylchloromethane under the conditions described for the preparation of trityl alkali lignin (13) and purified twice from chloroform-ether-light petroleum. Yield 1.0 g.

Found: C, 67.0; H, 6.0;  $\text{OCH}_3$ , 21.8%.

Calculated for  $\text{C}_{111}\text{H}_{118}\text{O}_{34}$  (1996): C, 66.8; H, 6.0;  $\text{OCH}_3$ , 21.8%.

### III. ACKNOWLEDGMENTS

The author is indebted to Messrs. Australian Paper Manufacturers Limited for permission to publish this work; and to Miss Betty Hickox for the semi-micro-analyses.

### IV. REFERENCES

- (1) HÄGGLUND, E., and KLINGSTEDT, F. W.—*Z. Physik Chem.* **152**: 295 (1931).
- (2) HERZOG, R. O., and HILLMER, A.—*Ber. deutsch. chem. Ges.* **64B**: 1288 (1931).
- (3) PEARL, I. A.—*J. Amer. Chem. Soc.* **64**: 1429 (1942).
- (4) HARRIS, E. E., and ADKINS, H.—*Paper Tr. J.* **107** (20): 38 (1938).
- (5) PHILLIPS, M.—*Science* **73**: 568 (1931).
- (6) DORÉE, C., and BARTON-WRIGHT, E. C.—*Biochem. J.* **21**: 290 (1927).
- (7) MARSHALL, H. B., BRAUNS, F. E., and HIBBERT, H.—*Canad. J. Res.* **13B**: 103 (1935).
- (8) POWELL, W. J., and WHITTAKER, H.—*J. Chem. Soc.* **127**: 132 (1925).
- (9) BECKMANN, E., LIESCHE, O., and LEHMANN, F.—*Biochem. Z.* **139**: 491 (1923).
- (10) PHILLIPS, M., and GOSS, M. J.—*J. Biol. Chem.* **125**: 241 (1938).
- (11) LAHEY, F. N., and MEREWETHER, J. W. T.—*Aust. J. Sci. Res. A* **1**: 112 (1948).
- (12) MEREWETHER, J. W. T.—*Ibid.* **1**: 241 (1948).
- (13) MEREWETHER, J. W. T.—*Ibid.* **2**: 117 (1949).

# THE CHEMICAL CONSTITUENTS OF AUSTRALIAN *FLINDERSIA* SPECIES\*

## II. BRAYLIN AND BRAYLEYANIN

By F. A. L. ANET†, G. K. HUGHES†, and E. RITCHIE†

[Manuscript received May 18, 1949]

### Summary

Two substances, *braylin*,  $C_{15}H_{14}O_4$ , and *brayleyanin*,  $C_{20}H_{24}O_4$ , have been isolated from the bark of *F. brayleyana* F. Muell. By degradation and synthesis they have been proved to be 6-methoxy-7, 8-(2',2'-dimethylpyrano-5',6')coumarin and 6-methoxy-7-( $\gamma\gamma$ -dimethylallyloxy)-8-( $\gamma\gamma$ -dimethylallyl)coumarin respectively.

## I. INTRODUCTION

*Flindersia brayleyana* F. Muell., a large umbrageous tree native to Queensland, has smooth, pale bark. A light petroleum extract of the bark gave, besides some oil, two crystalline substances which have been called *braylin* and *brayleyanin*. Both are soluble in the common organic solvents, insoluble in water, cold dilute acids and alkalis, are optically inactive, and give an orange-red colour with concentrated sulphuric acid. *Braylin* analyses for  $C_{15}H_{14}O_4$  and contains one methoxyl group. A carbon tetrachloride solution reacts with one mole of bromine to give a stable crystalline *dibromide*,  $C_{15}H_{14}O_4Br_2$ . The presence of a double bond is further demonstrated by the absorption of one mole of hydrogen in the presence of palladized charcoal. The *dihydrobraylin*,  $C_{15}H_{16}O_4$ , has the same crystalline habit as *braylin* and its melting point is not depressed by admixture with the latter substance. This is probably due to the formation of mixed crystals. *Dihydrobraylin*, unlike *braylin*, shows an intense blue fluorescence in alcohol. On addition of bromine in carbon tetrachloride solution *dihydrobraylin* reacts slowly to give an unstable *dibromide* which, on solution in alcohol, is converted to a stable *bromodihydrobraylin*,  $C_{15}H_{15}O_4Br$ . This indicates that *braylin* contains two double bonds, one of which has some aromatic character and this is confirmed by the uptake of two moles of hydrogen in the presence of Raney nickel with the formation of *tetrahydrobraylin*,  $C_{15}H_{18}O_4$ , which is more soluble in organic solvents than the less hydrogenated substances.

In Figure 1 the ultraviolet absorption spectra of *braylin*, *dihydrobraylin*,

\*This plant material was supplied by L. J. Webb, Research Officer, C.S.I.R.O., and is part of the survey being carried out by the C.S.I.R.O. and other institutions on the chemical constituents of the Australian flora.

†Chemistry Department, University of Sydney.

tetrahydrobraylin, and brayleyanin are shown and it is of interest to note the similarity of the spectra of the naturally occurring substances and the change in absorption brought about by reduction.

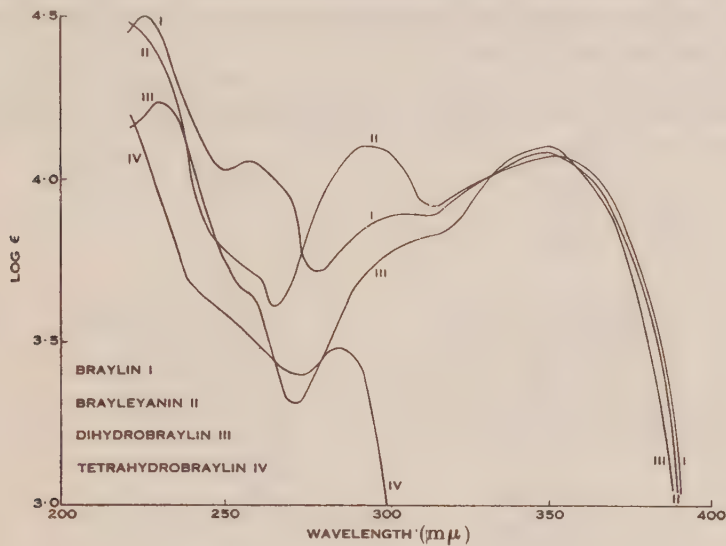


Fig. 1.—Ultraviolet absorption spectra of braylin, dihydrobraylin, tetrahydrobraylin, and brayleyanin. The measurements were made on ethyl alcoholic solutions using a Beckmann photoelectric quartz spectrophotometer.

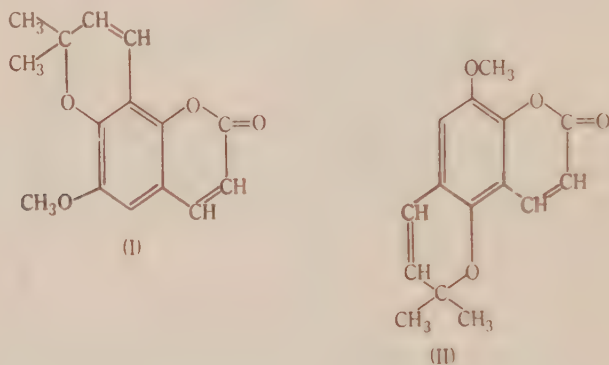
Both braylin and its dihydro derivative dissolve in hot aqueous sodium hydroxide to give yellow solutions from which they are precipitated unchanged by acids. However, tetrahydrobraylin gives a colourless solution when dissolved in hot alkali and *tetrahydrobraylinic acid*\*,  $C_{15}H_{20}O_5$ , is obtained on acidification. Vacuum sublimation reconverts this to tetrahydrobraylin.

These reactions suggest the presence of a coumarin structure for the natural product. This is further supported by the treatment of braylin and its dihydro and tetrahydro derivatives with dimethyl sulphate and sodium hydroxide to yield *methylbraylinic acid*,  $C_{16}H_{18}O_5$ , *methyldihydrobraylinic acid*,  $C_{16}H_{20}O_5$ , and *methyl-tetrahydrobraylinic acid*,  $C_{16}H_{22}O_5$ , respectively, each of which contains two methoxyl groups. The  $H_{18}$  and  $H_{20}$  acids are easily reduced to the  $H_{22}$  acid which is characteristic of cinnamic acids. Assuming the presence of a coumarin structure and one methoxyl group, an isoprene residue would account for the remaining five carbon atoms. This is indicated by the formation of  $\alpha$ -hydroxyisobutyric acid by permanganate oxidation of methylbraylinic acid. The function of one oxygen atom

\*The names assigned to this and other acids derived from braylin are in accord with those given to the acids obtained by the opening of the coumarin ring. Thus coumarin gives coumarinic acid, daphnetin gives daphnetinic acid etc.(1).

remains to be determined. A dimethylchromeno structure is probable since Robertson *et al.* (2, 3) have shown that substituted 2,2-dimethylbenzopyrans give  $\alpha$ -hydroxyisobutyric acid on oxidation. Bell and Robertson (4) have also demonstrated that such substances give acetone and a phenol on prolonged heating with aqueous alkalis. Under these conditions braylin also gives acetone and a phenol, m.p. 63°C., identical with the known 4-methoxyresorcinol.

The isolation of the above phenol definitely establishes the positions of the oxygen atoms attached to the benzene ring but the foregoing evidence still allows two structures for braylin.



On phytochemical grounds it appears probable that braylin is I: since all complex coumarins so far found in nature are related to simple, natural coumarins (6). Thus I can be said to be derived from 6,7-dihydroxycoumarin (aesculetin), which is widespread whereas II is related to 5,8-dihydroxycoumarin as yet unknown from natural sources.

Confirmation of this hypothesis is found in the easy cyclization of methyl-tetrahydrobraylinic acid (III) to a *ketone* (IV), a reaction not possible if the acid (V) is derived from II.

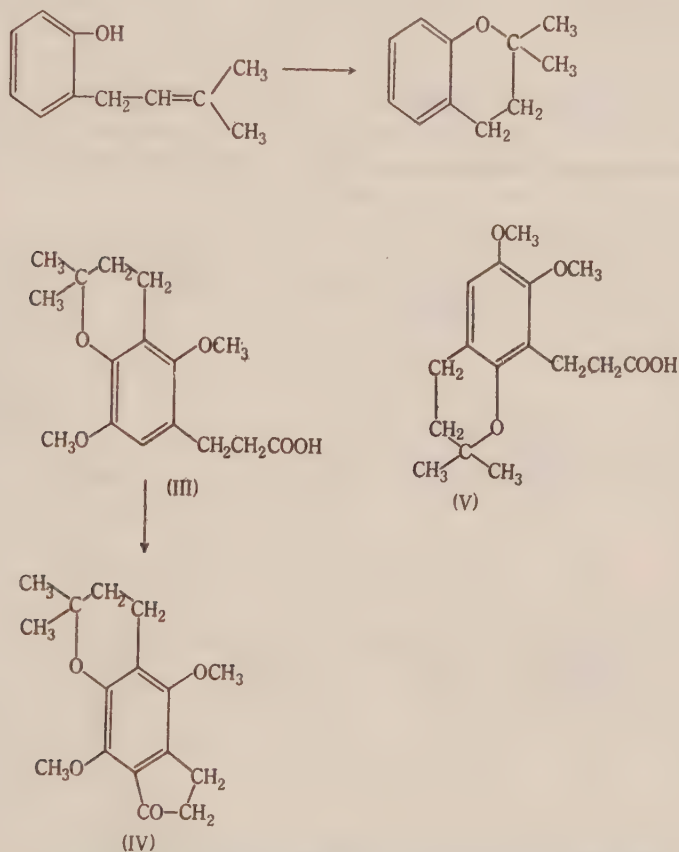
It is now necessary to discuss the chemistry of brayleyanin as it provided further proof of the structure assigned to braylin. Brayleyanin,  $C_{20}H_{24}O_4$ , contains one methoxyl group and readily absorbs bromine, but no crystalline derivative is obtained. On catalytic hydrogenation with Raney nickel in alcohol it affords a mixture of compounds one of which a *phenol*,  $C_{15}H_{16}O_4$  (VI) containing one methoxyl group, is easily isolated owing to its low solubility. This phenol is more readily obtained by warming brayleyanin with alcoholic hydrochloric acid for a short time. These two reactions suggest the presence of an allyloxy group and, as the residue lost is  $C_5H_8$ , it is most probably a  $\gamma\gamma$ -dimethylallyl ether. This is proved correct by the synthesis of brayleyanin from the sodium salt of the above phenol and  $\gamma\gamma$ -dimethylallyl bromide.

Both the phenol and brayleyanin on more prolonged treatment with hydro-



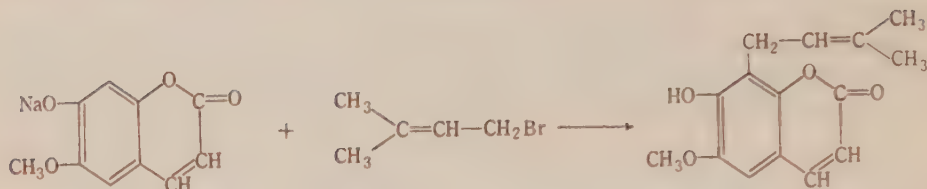
chloric acid in acetic acid give a non-phenolic product,  $C_{15}H_{16}O_4$ , identical with dihydrobraylin.

It has already been shown that dihydrobraylin is a dihydropyranocoumarin and therefore the above reactions are most readily explained by the cyclization of an *o*-( $\gamma\gamma$ -dimethylallyl) phenol to a chroman.

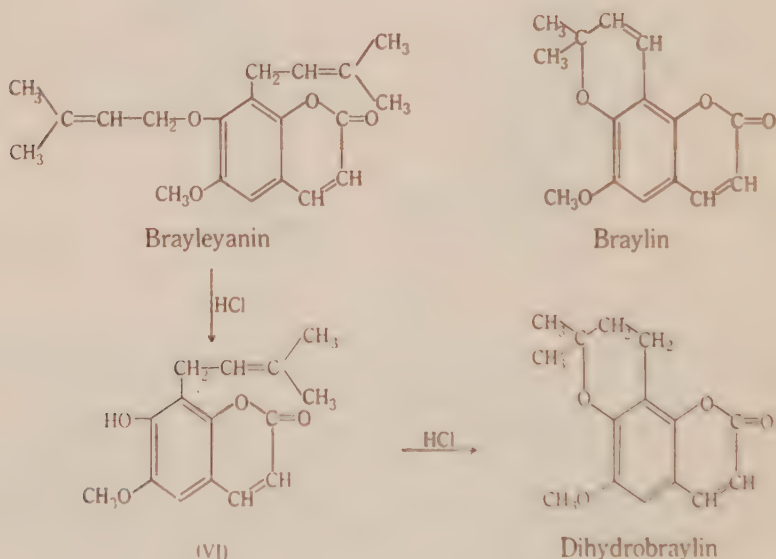


The possibility of the formation of a five-membered ring is excluded in view of the observation of Smith *et al.*(7) that although *o*-allyl phenols give coumarans, *o*-( $\gamma\gamma$ -dimethylallyl) phenols invariably give chromans.

The constitution of the phenol,  $C_{15}H_{16}O_4$ , is proved by its synthesis from the sodium salt of scopoletin (6-methoxy-7-hydroxycoumarin) by C-alkylation with  $\gamma\gamma$ -dimethylallyl bromide in benzene solution. The side chain must enter the 8 position. This reaction plus the preparation of brayleyanin previously described constitutes a total and unambiguous synthesis of brayleyanin and therefore of dihydrobraylin also.



Thus brayleyanin is 6-methoxy-7-( $\gamma\gamma$ -dimethylallyloxy)-8-( $\gamma\gamma$ -dimethylallyl)-coumarin and braylin is 6-methoxy-7,8-(2',2'-dimethylpyrano-5',6')coumarin. The latter is thus isomeric with xanthoxyletin(2, 8), alloxanthoxyletin(9), and luvange-  
tin(10).



It is interesting to note that while many complex coumarins theoretically derived from 1,2,3- and 1,3,5-trihydroxybenzenes are found in nature, braylin and brayleyanin are the first examples derived from 1,2,4-trihydroxybenzene.

## II. EXPERIMENTAL

(All melting points are uncorrected.)

The finely-milled bark (4.8 kg.) was continuously extracted with light petroleum until no further material was obtained on evaporation of the solvent. The combined liquors were concentrated to about 200 ml., allowed to cool, and the pale yellow crystals collected, recrystallized first from benzene-light petroleum and then from alcohol. Braylin (20 g.) was obtained in almost colourless flat rhombohedra, m.p. 150°C.

Found: C, 69.6; H, 5.5; MeO, 11.2, 11.5%; mol. wt. 240.

Calculated for  $C_{18}H_{14}O_4$ : C, 69.8; H, 5.5; 1MeO, 12.0%; mol. wt., 258.

By further concentration of the mother liquors of the original extract, crude brayleyanin

was obtained. After repeated recrystallization from light petroleum it formed colourless cubes, m.p. 95°C. (6 g.).

Found: C, 72.6; H, 7.5; MeO, 9.7%; mol. wt., 330.

Calculated for  $C_{20}H_{24}O_4$ : C, 73.2; H, 7.4; 1MeO, 9.5%; mol. wt., 328.

(a) *Braylin Dibromide*

Braylin (0.5 g.) was dissolved in the least amount of carbon tetrachloride and bromine added until the red colour was permanent. The dibromide, which precipitated, was recrystallized from carbon tetrachloride in small cubes, m.p. 196°C. (decomp.).

Found: C, 42.8; H, 3.5; Br, 39.3; MeO, 6.7%.

Calculated for  $C_{15}H_{14}O_4Br_2$ : C, 43.1; H, 3.4; Br, 38.3; 1MeO, 7.4%.

(b) *Dihydrobraylin*

Braylin (1 g.) in glacial acetic acid (80 ml.) was hydrogenated at room temperature in the presence of palladized charcoal (1 g. of 10%). Absorption of one mole of hydrogen was complete in 40 minutes.

After removal of the catalyst, the solvent was evaporated under reduced pressure, and the residue recrystallized from alcohol. Dihydrobraylin, large flat rhombohedra, m.p. 146°C., did not depress the melting point of braylin but gave an intense blue fluorescence in either alcohol or glacial acetic acid.

Found: C, 69.0; H, 6.2; MeO, 11.3%.

Calculated for  $C_{15}H_{16}O_4$ : C, 69.2; H, 6.2; 1MeO, 11.9%.

(c) *Bromodihydrobraylin*

To dihydrobraylin (0.2 g.) in carbon tetrachloride (5 ml.) was added bromine in the same solvent until a slight excess was present. The precipitate (m.p. about 130°C.) was recrystallized from alcohol as fine colourless needles, m.p. 160°C.

Found: Br, 23.6%.

Calculated for  $C_{15}H_{15}O_4Br$ : Br, 23.8%.

(d) *Tetrahydrobraylin*

Braylin (1 g.) in alcohol (40 ml.) was hydrogenated at room temperature in the presence of Raney nickel. Two moles were absorbed during several hours. Tetrahydrobraylin, after recrystallization from aqueous alcohol, was obtained as stout needles, m.p. 108°C. It showed no fluorescence in alcoholic solution.

Found: C, 68.2; H, 6.9; MeO, 11.4%.

Calculated for  $C_{15}H_{18}O_4$ : C, 68.7; H, 6.9; 1MeO, 11.8%.

(e) *Hydrolysis of Tetrahydrobraylin*

Tetrahydrobraylin (0.5 g.) was refluxed with aqueous sodium hydroxide (5 ml. of 5%) for 5 minutes. The precipitate obtained on acidifying the colourless solution was recrystallized from water, m.p. 149°C. This acid sublimed at 180°C. (bath temperature)/0.1 mm. and the product was identical (m.p. and mixed m.p.) with tetrahydrobraylin.

(f) *Methylbraylinic Acid*

To the deep orange solution of braylin (5 g.) in alcohol (40 ml.) and aqueous sodium hydroxide (50 ml. of 20%) was added, with shaking, dimethyl sulphate until the mixture was neutral. Sodium hydroxide solution (15 ml.) and a further amount of dimethyl sulphate were added in the same way. This was repeated. The solution was then made strongly alkaline and warmed on a water-bath for 15 minutes. The clear solution was then cooled, acidified, and the precipitate collected and recrystallized from methanol as colourless needles, m.p. 152°C. (5.5 g.). On repeated recrystallization the melting point was constant at 176°C. The change in melting point may be due to *cis-trans* isomerism; the lower melting compound was not analysed.

Found: C, 65.9; H, 6.2; MeO, 21.1%.

Calculated for  $C_{16}H_{18}O_5$ : C, 66.2; H, 6.2; 2MeO, 21.4%.

(g) *Methyldihydrobraylinic Acid*

Prepared as for methylbraylinic acid from dihydrobraylin and the product recrystallized from aqueous alcohol, m.p. 155°C.

(h) *Methyltetrahydrobraylinic Acid*

Prepared for methylbraylinic acid using either tetrahydrobraylin or tetrahydrobraylinic acid and recrystallized from benzene-light petroleum as colourless prisms, m.p. 102°C.

Found: C, 64.9; H, 7.7; MeO, 20.3%.

Calculated for  $C_{16}H_{22}O_8$ : C, 65.3; H, 7.5; 2MeO, 21.1%.

This acid was also prepared by the hydrogenation of either methylbraylinic acid or methyl-dihydrobraylinic acid in the presence of palladized charcoal. The theoretical amount of hydrogen was absorbed in 10 minutes and the yields were almost quantitative.

(i) *Oxidation of Methylbraylinic Acid*

Methylbraylinic acid (2.4 g.) was dissolved in hot acetone (300 ml.) and potassium permanganate (12 g.) added gradually. The brown precipitate was filtered and thoroughly extracted with caustic soda. The combined aqueous liquors were acidified and extracted with ether. Evaporation of the solvent left an oil which on distillation gave a fraction (b.p. 100°C./2 mm.) which solidified and after recrystallization from light petroleum had m.p. 78°C., undepressed by admixture with  $\alpha$ -hydroxyisobutyric acid, m.p. 79°C.

(j) *Fission of Braylin*

Braylin (1 g.) and caustic soda solution (50 ml. of 50%) were heated under reflux for 4 hours, water (50 ml.) added, and the mixture distilled until 20 ml. was collected. Aqueous dinitrophenylhydrazine hydrochloride was added to the distillate and the yellow precipitate collected, recrystallized from aqueous acetic acid, and then alcohol. Its m.p. 126°C. was undepressed by an authentic specimen of acetone dinitrophenylhydrazone.

The residual liquor from the distillation was acidified and filtered from a small amount of tar. It was then neutralized with sodium bicarbonate, saturated with ammonium sulphate, and extracted five times with ether. The ether solution was dried, the solvent removed, and the residue twice sublimed at reduced pressure. The colourless crystals had m.p. 63°C., undepressed on admixture with 4-methoxyresorcinol. The *p*-nitrobenzoate of the phenol was made using *p*-nitrobenzoyl chloride in pyridine and recrystallized from methanol as plates, m.p. 151-152°C., undepressed by 4-methoxyresorcinol di-*p*-nitrobenzoate.

Found: N, 6.4%.

Calculated for  $C_{21}H_{14}O_9N_2$ : N, 6.4%.

(k) *Cyclization of Methyltetrahydrobraylinic Acid*

The acid was warmed with five times its weight of sulphuric acid on a water-bath for 2 minutes, poured into water, and mixed with an alcoholic solution of dinitrophenylhydrazine. The precipitate was collected and recrystallized from ethyl acetate, m.p. 260°C.

Found: N, 12.1%.

Calculated for  $C_{22}H_{24}O_7N_4$ : N, 12.3%.

(l) *Hydrogenation of Brayleyanin*

This was carried out in the usual way in alcohol with Raney nickel as catalyst. Absorption was slow after 2.5 moles had been taken up. After filtration and concentration to a small volume, crystals separated which on recrystallization from alcohol had m.p. 171-172°C.

Found: C, 68.9; H, 6.4; MeO, 11.4%.

Calculated for  $C_{18}H_{16}O_4$ : C, 69.2; H, 6.2; 1MeO, 11.9%.

After removal of solvent from the mother liquors only oils were obtained and no other pure compound was isolated.



(m) *Hydrolysis of Brayleyanin*

Brayleyanin (1 g.) and hydrochloric acid (1 ml.) in methanol (20 ml.) were warmed for 5 minutes on a water-bath, poured into water, and the precipitated phenol recrystallized from alcohol. The product (0.67 g.) had m.p. 171-172°C., undepressed on admixture with the product obtained by hydrogenation of brayleyanin.

(n) *6-Methoxy-7-Hydroxy-8-( $\gamma\gamma$ -Dimethylallyl)coumarin*

The scopoletin used was prepared according to the method of Head and Robertson(11) and had m.p. 204°C. The vacuum-dried yellow sodium salt of scopoletin, from scopoletin (0.154 g.) and sodium (0.02 g.), was refluxed with  $\gamma\gamma$ -dimethylallyl bromide (0.2 ml.) in benzene (5 ml.) for 8 hours, diluted with ether, and extracted with sodium hydroxide solution. The aqueous layer was acidified, extracted with ether, the solvent removed, and the residue recrystallized from alcohol. The product (0.11 g.) had m.p. 170-172°C., undepressed on admixture with the product obtained by hydrolysis of brayleyanin.

(o) *Synthesis of Brayleyanin*

6-Methoxy-7-hydroxy-8-( $\gamma\gamma$ -dimethylallyl)coumarin (0.26 g.),  $\gamma\gamma$ -dimethylallyl bromide (0.16 g.), and sodium ethoxide (from 0.023 g. sodium) in alcohol (20 ml.) were refluxed for 8 hours. After cooling, dilute sodium hydroxide was added, and the mixture extracted with ether. The solvent was removed from the dried organic layer and the residue recrystallized from light petroleum had m.p. 95°C., undepressed on admixture with brayleyanin.

(p) *Conversion of Brayleyanin to Dihydrobraylin*

Brayleyanin (0.5 g.) was heated with glacial acetic acid (20 ml.) and hydrochloric acid (1 ml.) for 3 hours. The solvent was removed under reduced pressure and the residue recrystallized twice from alcohol (0.25 g.) had m.p. 145-146°C., undepressed on admixture with dihydrobraylin.

### III. ACKNOWLEDGMENTS

The authors are greatly indebted to Mr. L. J. Webb, C.S.I.R.O., for collection of the bark and for general advice; and to Miss J. Fildes, microanalyst, for all the analyses.

### IV. REFERENCES

- (1) SETHNA, S., and SHAH, N.—*Chem. Rev.* **36**: 1 (1945).
- (2) BELL, J., ROBERTSON, A., and SUBRAMANIAM, T.—*J. Chem. Soc.* **1936**: 627 (1936).
- (3) BELL, J., BRIDGE, W., and ROBERTSON, A.—*Ibid.* **1937**: 1542 (1937).
- (4) BELL, J., and ROBERTSON, A.—*Ibid.* **1936**: 1828 (1936).
- (5) SEKA, R., and KALLIR, P.—*Ber. dtsh. chem. Ges.* **64**: 909 (1931).
- (6) SPÄTH, E.—*Ibid.* **70**: 83 (1937).
- (7) SMITH, L. I., UNGNADE, H. E., HOEHN, H. E., and WAWZONEK, S.—*J. Org. Chem.* **4**: 305, 311 (1939).
- (8) ROBERTSON, A., and SUBRAMANIAM, T.—*J. Chem. Soc.* **1937**: 286 (1937).
- (9) ROBERTSON, A., and SUBRAMANIAM, T.—*Ibid.* **1937**: 1545 (1937).
- (10) SPÄTH, E., and SCHMID, H.—*Ber. dtsh. chem. Ges.* **74**: 193 (1941).
- (11) HEAD, F., and ROBERTSON, A.—*J. Chem. Soc.* **1931**: 1241 (1931).

# A SYNTHESIS OF HYGRINE AND OF CUSCOHYGRINE\*

By E. F. L. J. ANET†, G. K. HUGHES†, and E. RITCHIE†

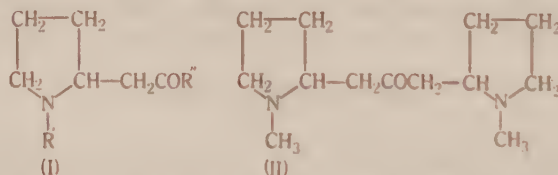
[Manuscript received August 31, 1949]

## Summary

The synthesis of hygrine and of cuscohygrine from 4-methylaminobutanal is described, and the stereo-chemistry of cuscohygrine is discussed.

## I. INTRODUCTION

In a classical paper published in 1917 Robinson(1) proposed a theory of the biogenesis of several of the important groups of alkaloids. This theory became widely known and has directly stimulated attempts to synthesize alkaloids under "physiological conditions," i.e. by the use of mild reagents, in dilute aqueous solution close to neutrality at or near room temperature. The brilliant successes achieved mainly by the schools of Robinson, Schöpf, and Hahn (Henry 2) have provided a strong argument in support of the essential validity of Robinson's theory.



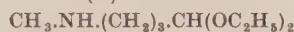
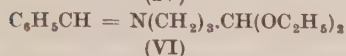
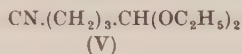
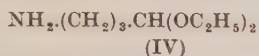
According to the theory hygrine (I,  $R' = R'' = \text{CH}_3$ ) and cuscohygrine (II) are formed by the condensation of acetonedicarboxylic acid with one and two molecular proportions respectively, of 4-methylaminobutanal (or its ring tautomer), but for nearly twenty years no attempt to confirm these ideas appears to have been made. Then in 1936 Robinson(3) announced that a synthesis of hygrine on these lines had been accomplished but no detailed account of this work has since appeared. A year later Schöpf(4) reported an unsuccessful attempt to prepare *norhygrine* (I,  $R' = \text{H}$ ,  $R'' = \text{CH}_3$ ) from 4-aminobutanal and acetoacetic acid, although he found that the same aldehyde and benzoylacetic acid gave (I,  $R' = \text{H}$ ,  $R'' = \text{C}_6\text{H}_5$ ). However, two syntheses of hygrine(5,6) and one of cuscohygrine(7) using the conventional "brutal" methods have been recorded.

It has now been found that both hygrine and cuscohygrine can be synthesized under "physiological conditions" in good yields from 4-methylaminobutanal. This aldehyde in the form of its stable *diethyl-acetal* (III) was prepared by two methods from 4-amino-1,1-diethoxybutane (IV)(8). By the action of chloroform and potassium hydroxide, IV was converted to 4-isocyano-1,1-diethoxybutane (V). The yields

\*A brief account of this synthesis appeared in *Nature* **163**: 289 (1949).

†Chemistry Department University of Sydney.

were poor even when the usually excellent procedure of Malatesta(9) was employed, but much unused IV could be recovered. Reduction of V by sodium and alcohol then furnished III in good yield. The second method based on the well-known Decker and Becker reaction was more satisfactory. Benzaldehyde and IV readily gave



*benzylidene 4-amino-1,1-diethoxybutane* (VI) which on treatment with methyl iodide gave a quaternary salt and this was very rapidly hydrolysed to III and benzaldehyde, simply by dissolving in water.

When a dilute faintly acid solution of III was warmed for a short time, then cooled and neutralized, and added to a dilute solution of acetoacetic acid buffered to pH 7, a base,  $\text{C}_8\text{H}_{15}\text{ON}$ , was produced in about 70 per cent. yield. Although it has not been possible to compare it with an authentic specimen, the following comparisons of melting points and boiling points show that there can be no doubt that the synthetic product is hygrine.

	<i>Synthetic Product</i>	<i>Natural Hygrine</i>
Base, b.p.	104°C./40 mm.	92-95°C./20 mm. <sup>(a)</sup> 111-113°C./50 mm. <sup>(a)</sup> 79-81°C./11-12 mm. <sup>(b)</sup>
Oxime, m.p.	124°C.	116-120°C. <sup>(c)</sup> 125°C. <sup>(b)</sup>
Picrate, m.p.	148°C.	148°C. <sup>(d)</sup> 158°C. <sup>(b)</sup>
Oxime picrate, m.p.	160°C.	160°C. <sup>(c)</sup> 160°C. <sup>(b)</sup>
<sup>(a)</sup> Liebermann and Cybulski(10).		<sup>(c)</sup> Liebermann and Kühling(11).
<sup>(b)</sup> Hess(5).		<sup>(d)</sup> Liebermann(12).

Also, since Späth and Kittel(13) have shown that natural hygrine is optically inactive, the previously recorded small rotations(10,5) being due to the presence of hygroline, the identity of the natural with the synthetic product is complete.

In a similar manner III was condensed with benzoylacetic acid to yield the *phenyl-analogue* (I,  $\text{R}' = \text{CH}_3$ ,  $\text{R}'' = \text{C}_6\text{H}_5$ ) and with nicotinylacetic acid to the  *$\beta$ -pyridyl-analogue* (I,  $\text{R}' = \text{CH}_3$ ,  $\text{R}'' = \beta\text{-C}_5\text{H}_4\text{N}$ ) of hygrine.

The reaction between III and acetonedicarboxylic acid afforded cuscohygrine, together with hygrine. The proportion of the two bases varied with the pH of the reaction mixture as shown in the table.

	pH	5	7	9.2
Yield of cuscohygrine (per cent.)		15	54	38
Yield of hygrine (per cent.)		26	9	7

The cuscohygrine was identified by direct comparison of several of its derivatives with specimens prepared from authentic natural cuscohygrine. The melting

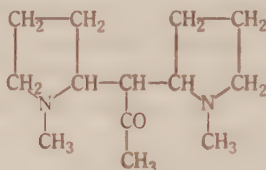
points observed were as follows:

	<i>Synthetic</i>	<i>Natural</i>	<i>Mixed</i>
Dinitrate	209°C.	209°C. <sup>(a)</sup>	209°C.
Dipicrate	213-214°C.	213-214°C. <sup>(b)</sup>	213-214°C.
Dimethiodide	244°C.	244°C. <sup>(a)</sup>	244°C.

(a) The same values are given by Hess and Bappert(14).

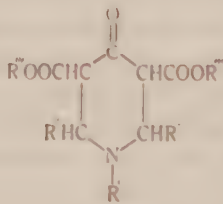
(b) Hess and Fink(15) give 215°C.

Before the structure of cuscohygrine was definitely settled by synthesis(7) an alternative formula (VII) suggested by Hess and Fink(15) received some consideration. It was therefore of interest to attempt the synthesis of VII by the action of excess III on acetoacetic acid. The only pure substance that could be isolated, however, from a reaction conducted at pH 7 was hygrine.

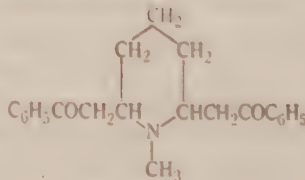


(VII)

One further point concerning the structure of cuscohygrine must be discussed. In formula II there are two similar asymmetric carbon atoms and therefore cuscohygrine may be either the *dl* or the *meso* form. Neither the natural nor the synthetic product, which it should be emphasized are identical, can be a mixture of the two forms since each behaves as a pure chemical individual giving pure crystalline derivatives in high yields. The resolution of natural cuscohygrine, which is optically inactive(10) was unsuccessfully attempted by Hess and Anselm(16) and attempts now made to resolve the synthetic base through its *d*-tartrate, *d*-antimonyltartrate or *d*-2-bromo-8-camphorsulphonate were abortive because the salts could not be crystallized. Although these experiments prove nothing, there is some positive evidence that cuscohygrine is the *meso* form. Mannich(17) has shown that in the



(VIII)



(IX)

$\gamma$ -piperidones (VIII), produced by the reaction between an aliphatic aldehyde, a primary aliphatic amine, and an ester of acetonedicarboxylic acid, the alkyl groups



in the 2 and 6 positions are always *cis*, i.e. that carbon atoms 2 and 6 are asymmetric but similar. Again, Schöpf and Lehmann(18), by condensing glutaraldehyde with methylamine and benzoylacetic acid at pH 4 to 5, obtained high yields of lobelanine (IX), in which the phenacyl residues are *cis* and which is a *meso* form. Both of these reactions and the synthesis of cuscohygrine are essentially reactions between an aliphatic amine, an aliphatic aldehyde, and a reactive methylene group, i.e. they are variations of the Mannich reaction. If it be assumed that the same steric features are present in all three variations, it follows that cuscohygrine is the *meso* form.

## II. EXPERIMENTAL

### (a) 4-*iso*Cyano-1, 1-Diethoxybutane (V)

(i) To a stirred suspension of powdered potassium hydroxide (80 g.) in pure dry benzene (200 ml.) was gradually added a solution of 4-amino-1,1-diethoxybutane (30 g.) in chloroform (75 g.). A vigorous reaction soon commenced and the addition was regulated so that the mixture refluxed gently. After a further 15 minutes refluxing the mixture was cooled, filtered, and the residue washed with benzene. Distillation of the combined filtrate and washings under reduced pressure gave eventually unused acetal (15 - 18 g.; b.p. up to 80°C./1.5 mm.) and the product, a colourless oil (3.2 - 4.8 g.; 10 - 15%) b.p. 80°C./1.5 mm.

(ii) Powdered potassium hydroxide (20 g.) was added in four portions during 1 hour to a refluxing solution of 4-amino-1,1-diethoxybutane (12 g.) in chloroform (20 ml.). After a further 10 minutes the mixture was cooled, filtered, and the residue washed with chloroform. Distillation under reduced pressure yielded unused acetal (6.7 g.) and the product (2.3 g.; 15 - 25%).

Found: N, 8.2%.

Calculated for  $C_9H_{17}O_2N$ : N, 8.2%.

### (b) Benzylidene 4-Amino-1,1-Diethoxybutane (VI)

4-Amino-1,1-diethoxybutane (16 g.; 1 mol.), freshly distilled benzaldehyde (11.7 g.; 1.1 mol.), and anhydrous potassium carbonate (2.8 g.; 0.2 mol.) were allowed to stand overnight. The supernatant liquid was decanted from the sludge which was washed with ether. Distillation of the combined liquid and washings gave a colourless viscous oil (22.5 g.; 90%), b.p. 155°C./3 mm.

Found: C, 71.6; H, 9.4; N, 5.6%.

Calculated for  $C_{15}H_{22}O_2N$ : C, 72.0; H, 9.3; N, 5.6%.

### (c) 4-Methylamino-1,1-Diethoxybutane (III)

(i) Sodium (3.8 g.) was added in one lot to a solution of the *isonitrile* (4.3 g.) in anhydrous alcohol (60 ml.). When reaction had ceased water (10 ml.) was added and as much alcohol as possible removed on the water-bath. The residue was extracted five times with ether, the extracts dried and distilled. The product (3.7 g.; 85%) was a colourless liquid, b.p. 95°C./12 mm. completely miscible with water.

(ii) A mixture of pure dry methyl iodide (28 g.; 2 mol.) and VI (25 g.; 1 mol.) was allowed to stand at room temperature for 3 days. The quaternary salt which had separated as a dark red very viscous lower layer, was washed with dry ether and then shaken with a small volume of water. In less than 2 minutes it had dissolved and had been completely hydrolysed. The mixture was quickly shaken with ether to remove benzaldehyde, then immediately saturated with potassium carbonate and extracted five times with ether. The product (12 - 14 g.; 70 - 80%) was isolated by distillation under reduced pressure.

Found: N, 8.0%.

Calculated for  $C_8H_{21}O_2N$ : N, 8.0%.

(d) *Hygrine* ( $I$ ,  $R' = R'' = CH_3$ )

4-Methylamino-1,1-diethoxybutane (3.5 g.; 1 mol.) was made acid to Congo red with hydrochloric acid (1%), warmed on the water-bath for 10 minutes, cooled, and quickly neutralized with dilute potassium hydroxide solution. This solution of 4-methylaminobutanal was immediately added to a neutralized solution of acetoacetic acid, prepared from ethyl acetoacetate (5.2 g.; 2 mol.) and phosphate buffer (100 ml. of M/10) of pH 7. After making up the volume to 300 ml. the mixture was allowed to stand overnight at room temperature. Then after saturating with potassium carbonate, it was extracted seven times with ether. The hygrine (2.0 g.; 70%) isolated by distillation under reduced pressure had b.p.  $104^\circ\text{C}$ . 40 mm. Since it rapidly absorbed carbon dioxide from the air it was analysed in the form of the derivatives described below.

The *oxime* was prepared by heating hygrine (0.85 g.), hydroxylamine hydrochloride (0.42 g.), potassium hydroxide (0.36 g.), and water (7 ml.) on the water-bath for  $1\frac{1}{2}$  hours. After cooling, the crystalline precipitate was collected, washed with a little water, dried and recrystallized from petroleum (b.p.  $60\text{--}90^\circ\text{C}$ .) from which it separated as colourless needles, m.p.  $124^\circ\text{C}$ .

Found: C, 61.4; H, 10.3; N, 17.8%.

Calculated for  $C_8H_{16}ON_2$ : C, 61.5; H, 10.3; N, 17.9%.

The *oxime picrate* crystallized from ethanol in yellow needles, m.p.  $160^\circ\text{C}$ .

Found: N, 18.0%.

Calculated for  $C_{14}H_{19}O_8N_5$ : N, 18.2%.

The *picrate* formed deep yellow prisms from ethanol, m.p.  $148^\circ\text{C}$ .

Found: N, 15.1%.

Calculated for  $C_{14}H_{18}O_8N_4$ : N, 15.1%.

(e) (*N*-Methyl- $\alpha$ -Pyrrolidyl) Methylphenylketone ( $I$ ,  $R' = CH_3$ ,  $R'' = C_6H_5$ )

By the same procedure as that used for hygrine, ethyl benzoylacetate (4 g.) was condensed with 4-methylamino-1,1-diethoxybutane (3 g.). The product (2.7 g.; 80%) which had b.p.  $150^\circ\text{C}$ . 3 mm. was a colourless viscous liquid which absorbed carbon dioxide and turned brown on exposure to air.

The *nitrate* precipitated by the slow addition of concentrated nitric acid to an ethereal solution of the base, crystallized from ethanol in colourless prisms, m.p.  $123\text{--}124^\circ\text{C}$ .

Found: C, 58.2; H, 6.9%.

Calculated for  $C_{13}H_{18}O_4N_2$ : C, 58.6; H, 6.8%.

(f) (*N*-Methyl- $\alpha$ -Pyrrolidyl) Methyl- $\beta$ -Pyrrolidylketone ( $I$ ,  $R' = CH_3$ ,  $R'' = \beta\text{-C}_4\text{H}_4\text{N}$ )

The base (1.4 g.; 40%) prepared from ethyl nicotinyacetate (4 g.) and 4-methylamino-1,1-diethoxybutane (3 g.) was a colourless viscous liquid, b.p.  $163^\circ\text{C}$ . 1.3 mm., which quickly darkened on exposure to air.

The *dipicrate* crystallized from acetone in yellow prisms, m.p.  $164\text{--}165^\circ\text{C}$ .

Found: N, 16.7%.

Calculated for  $C_{21}H_{22}O_{13}N_8$ : N, 16.9%.

(g) *Euscobygrine* (II)

A neutralized solution of 4-methylaminobutanal prepared from the acetal (2.8 g.; 2 mol.) was added to a neutralized solution of acetonedicarboxylic acid (1.17 g.; 1 mol.) mixed with phosphate buffer (110 ml. of M/10) of pH 7, then the volume made up to 250 ml., and the mixture allowed to stand overnight at room temperature. After saturating with potassium carbonate it was extracted with ether seven times, the extracts dried and the ether removed. The residue on fractionation under reduced pressure gave hygrine (0.2 g.; 9%), b.p. up to  $120^\circ\text{C}$ . 2 mm., and euscobygrine (0.97 g.; 54%), b.p.  $120^\circ\text{C}$ . 2 mm. The base which absorbed carbon dioxide from the air, was analysed as the derivatives below.

The *dimethiodide* prepared by adding excess methyl iodide to an alcoholic solution of the base, crystallized from ethanol in small colourless prismatic needles, m.p.  $244^\circ\text{C}$ .

Found : C, 35.2 ; H, 6.0 ; N, 5.5 ; I, 50.4%.

Calculated for  $C_{18}H_{30}ON_2I_2$  : C, 35.4 ; H, 6.0 ; N, 5.5 ; I, 50.0%.

The *dinitrate*, precipitated by slowly adding concentrated nitric acid to an ethereal solution of the base, crystallized from ethanol in short colourless needles, m.p. 209°C. (decomp.).

Found : N, 15.9%.

Calculated for  $C_{18}H_{30}O_7N_4$  : N, 16.0%.

The *dipicrate* crystallized from acetone in pale yellow needles, m.p. 213-214°C.

Found : N, 16.3%.

Calculated for  $C_{25}H_{30}O_{15}N_8$  : N, 16.4%.

(g) *Attempted Synthesis of VII*

The reaction between ethyl acetoacetate (2.6 g. ; 1 mol.) and 4-methylamino-1,1-diethoxybutane (7 g. ; 2 mol.) under the conditions described above gave hygrine (1.5 g. ; 54%), b.p. 110°C./50 mm. No material of higher boiling point could be isolated.

### III. ACKNOWLEDGMENTS

The authors are grateful to Miss J. Fildes for the analyses, to Professor R. L. Shriner, State University of Iowa, who kindly supplied an authentic specimen of cuscohygrine and to the Commonwealth Research Grant Committee of the University of Sydney for a scholarship awarded to one of them (E.F.L.J.A.).

### IV. REFERENCES

- (1) ROBINSON, R.—*J. Chem. Soc.* **111** : 876 (1917).
- (2) HENRY, T. A.—“The Plant Alkaloids.” 4th Ed. (Churchill : London, 1949).
- (3) ROBINSON, R.—*J. Chem. Soc.* **1936** : 1079 (1936).
- (4) SCHÖPF, C.—*Z. angew. Chem.* **50** : 797 (1937).
- (5) HESS, K.—*Ber. deutsch. chem. Ges.* **46** : 4104 (1913).
- (6) SORM, F.—*Chem. Abstr.* **42** : 558 (1948).
- (7) SPÄTH, E., and TUPPY, H.—*Ibid.* **42** : 8795 (1948).
- (8) MANSKE, R. H. F.—*Canad. J. Res.* **5** : 592 (1931).
- (9) MALATESTA, L.—*Gazz. Chim. Ital.* **77** : 238 (1947).
- (10) LIEBERMANN, C., and CYBULSKI, G.—*Ber. deutsch. chem. Ges.* **28** : 578 (1895).
- (11) LIEBERMANN, C., and KÜHLING, O.—*Ibid.* **26** : 851 (1893).
- (12) LIEBERMANN, C.—*Ibid.* **22** : 675 (1889).
- (13) SPÄTH, E., and KITTEL, F.—*Ibid.* **76** : 942 (1943).
- (14) HESS, K., and BAPPERT, R.—*Liebigs Ann.* **441** : 137 (1925).
- (15) HESS, K., and FINK, H.—*Ber. deutsch. chem. Ges.* **53** : 781 (1920).
- (16) HESS, K., and ANSELM, F.—*Ibid.* **54** : 2310 (1921).
- (17) MANNICH, C.—*Arch. Pharm. Berl.* **272** : 323 (1934).
- (18) SCHÖPF, C., and LEHMANN, G.—*Liebigs Ann.* **518** : 1 (1935).



# ALKALOIDS OF THE AUSTRALIAN RUTACEAE: *ACRONYCHIA BAUERI*

## II. SOME REACTIONS OF THE ALKALOID ACRONYCINE

By R. D. BROWN, L. J. DRUMMOND, F. N. LAHEY\*, and W. C. THOMAS

[Manuscript received August 5, 1949]

### Summary

Acronycine,  $C_{20}H_{19}O_3N$ , is readily converted to noracronycine,  $C_{19}H_{17}O_3N$ , a weak base which contains no methoxyl. Acronycine contains a reactive double bond, shown to be present in a dimethylpyran ring by oxidation to an acid which, on heating, yields  $\alpha$ -hydroxyisobutyric acid together with a phenol,  $C_{14}H_{11}O_3N$ , and its methyl ether,  $C_{15}H_{13}O_3N$ . The action of alcoholic potash on acronycine results in the formation of two phenols, both of which are readily reconverted to acronycine. The bromination of acronycine has also been studied.

## I. INTRODUCTION

In Part I (1) the isolation of a new alkaloid, acronycine was recorded. Elementary analyses showed it to have the molecular formula  $C_{20}H_{19}O_3N$  and to contain one methoxyl and one methylimino group while negative results were obtained from the common reactions for hydroxyl and carbonyl groups.

Like the alkaloids of the related *Melicope fareana* examined by Crow and Price (2) acronycine is susceptible to hot mineral acids although it is not affected so readily. Hot alcoholic hydrochloric acid yields a polymeric amorphous product, but heating acronycine hydrochloride in the dry state brings about demethylation with the formation of noracronycine,  $C_{19}H_{17}O_3N$ , a bright yellow crystalline solid, m.p. 200.5-201°C. norAcronycine is a much weaker base than acronycine and does not form stable salts. It contains one methylimino group but no methoxyl. It shows no phenolic properties and is not methylated by diazomethane. However, on treatment in acetone solution with dimethyl sulphate and potassium carbonate it is converted back to acronycine. With acetic anhydride it gives a *monoacetate*.

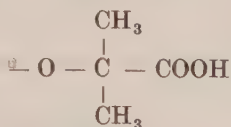
Acronycine contains a reactive double bond, catalytic reduction in the presence of Raney nickel giving *dihydroacronycine*, m.p. 140-141.5°C. Hydrochloric or hydrobromic acid converts dihydroacronycine to *norDihydroacronycine* without the polymerization which was experienced with acronycine. *norDihydroacronycine* may also be prepared by heating the orange hydrochloride of dihydroacronycine in the dry state.

Oxidation of acronycine with potassium permanganate in acetone solution gives a *dibasic acid*,  $C_{20}H_{19}O_7N$ , without loss of carbon so that the double bond is present in a ring system. This dibasic acid very readily loses carbon dioxide to yield a monobasic acid,  $C_{19}H_{19}O_5N$ , which for convenience we have named *acronycinic*

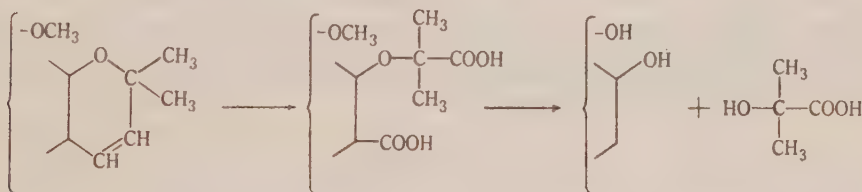
\*Present address: Chemistry Department, University of Queensland.



*acid*. Thermal decomposition of acronycinic acid above its melting point (227°C.) yields three products, a phenol,  $C_{14}H_{11}O_3N$ , m.p. 286-289°C., a weak base,  $C_{15}H_{13}O_3N$ , m.p. 174-175°C., which has been shown to be the methyl ether of the phenol, and a volatile acid identified as  $\alpha$ -hydroxyisobutyric acid. The formation of the last named indicates the presence of the group



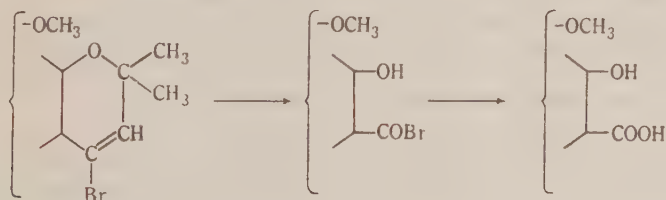
in acronycinic acid which in turn must result from the oxidation of a dimethylpyran ring in the original alkaloid. The phenol,  $C_{14}H_{11}O_3N$ , can also be obtained by the action of hydrobromic acid on acronycinic acid or on the base,  $C_{15}H_{13}O_3N$ , above. On treatment with diazomethane the phenol,  $C_{14}H_{11}O_3N$ , is converted into the base,  $C_{15}H_{13}O_3N$ , while further treatment with dimethyl sulphate gives a product,  $C_{16}H_{15}O_3N$ , containing two methoxyl groups. The formation of the phenol,  $C_{14}H_{11}O_3N$ , can be represented by the following partial formulae:



*nor*Acronycine on oxidation yields *noracronycinic acid* which on heating gives only the phenol,  $C_{14}H_{11}O_3N$ , and  $\alpha$ -hydroxyisobutyric acid.

Acronycine reacts with bromine in chloroform solution, addition to the double bond apparently occurring, followed by the rapid elimination of hydrogen bromide, and the precipitation of the red hydrobromide of *monobromacronycine*, m.p. 195°C. The bromine is undoubtedly attached to one of the carbon atoms joined by the double bond of the pyran ring for on oxidation acronycinic acid is formed.

Ozonolysis of bromacronycine gives a phenolic acid (I) while acronycine gives a phenolic aldehyde (see Part III of this series). This suggests that the bromine is located on the carbon atom adjacent to the aromatic ring.

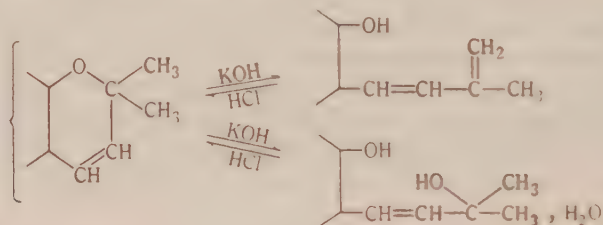


(I)

Excess bromine on bromacronycine introduces one more bromine atom and causes demethylation, to give *nordibromacronycine*, a substance capable of existing

in dimorphic forms, one yellow and the other red. The same product is obtained by the bromination of *nora*acronycine in two stages as above.

Concentrated potash solution on naturally occurring phenols which contain a dimethylbenzpyran ring such as toxicarol and deguelin, causes the opening of the pyran ring and the elimination of acetone(3, 4, 5). Alcoholic potash on acronycine does not produce acetone although it is thought that the pyran ring is opened as two phenols are produced, one of them isomeric with acronycine and the other analysing for a dihydrate of acronycine. Both of these phenols are readily converted to acronycine by the action of warm hydrochloric acid. In view of the known effect of alkali on the benzpyran nucleus it is probable that the action of alcoholic potash on acronycine can be represented as follows:



With the first of these two, where an extra conjugated double bond is introduced into the structure, spectrographic evidence to support this has been obtained which will be submitted in a separate paper.

## II. EXPERIMENTAL

All melting points are corrected unless otherwise stated. Microanalyses are by Dr. C. Tettweiler and Mr. N. L. Lottkowitz.

### (a) *nor*Acronycine

(i) Acronycine (2 g.) was dissolved in alcohol (30 ml.) and to this solution was added concentrated hydrochloric acid (5 ml.). The deep red solution was heated under reflux on a water-bath for 5 to 6 hours. An orange-yellow solid separated which was filtered off. All attempts to crystallize this yielded only an amorphous powder of indefinite melting point. It was not further examined.

(ii) Acronycine hydrochloride was formed by treating acronycine with 10% hydrochloric acid, heating to effect solution, and allowing to cool, when the hydrochloride separated in bright red needles which were dried over calcium chloride. Acronycine hydrochloride (0.5 g.) was heated in a test tube to 130°C. until effervescence ceased. A volatile substance, presumably methyl chloride, was eliminated and a reddish liquid, which set to a glass on cooling, remained. This was dissolved in chloroform (5 ml.), alcohol (10 ml.) added, and most of the chloroform distilled off in a water-bath. While still hot, crystallization of *nor*acronycine began, giving bright yellow needles, m.p. 200.5-201°C. Yield 90%.

Found: C, 74.1; H, 5.6; N, 4.5;  $\text{CH}_3\text{O}$ , nil;  $\text{CH}_3\text{N}$ , 8.5%.

Calculated for  $\text{C}_{19}\text{H}_{17}\text{O}_3\text{N}$ : C, 74.3; H, 5.5; N, 4.6;  $\text{CH}_3\text{O}$ , nil;  $\text{CH}_3\text{N}$ , 9.4%.

*nor*Acronycine can also be crystallized from acetic acid or ethyl acetate. It is readily soluble in chloroform and benzene, and sparingly soluble in light petroleum. It is a much weaker base than acronycine, being insoluble in 10% hydrochloric acid but soluble in concentrated hydrochloric acid. It does not form an insoluble hydrochloride as does acronycine. It is insoluble in aqueous alkali.

(iii) *Methylation of norAcronycine*.—To a solution of *noracronycine* (0.3 g.) in acetone (20 ml.) was added anhydrous potassium carbonate (4 g.) and dimethyl sulphate (2 ml.) and the whole heated under reflux for 36 hours, fresh potassium carbonate and dimethyl sulphate being added from time to time. The red solution was filtered and concentrated under reduced pressure, when a salt, presumably the methosulphate of acronycine, separated. This was filtered and treated with ammonia. Recrystallization of the resulting base yielded acronycine, m.p. 175°C.

(iv) *Acetyl nor acronycine*.—*norAcronycine* (0.2 g.) was acetylated by heating with acetic anhydride (5 ml.) and a few drops of pyridine for 3 hours. On cooling, the *acetyl derivative* crystallized in short needles. It could be recrystallized from acetic anhydride or dry benzene, separating in yellow needles, m.p. 166-166.5°C. Recrystallization from alcohol brought about partial alcoholysis to *noracronycine*.

Found: C, 72.0; H, 5.6; N, 4.0%.

Calculated for  $C_{21}H_{19}O_4N$ : C, 72.2; H, 5.4; N, 4.0%.

#### (b) *Dihydroacronycine*

The alkaloid (1 g.) dissolved in alcohol (50 ml.) was shaken with hydrogen in the presence of Raney nickel. After absorption of hydrogen was complete (equivalent to that required for one double bond) the solution was filtered. This solution had a strong blue fluorescence in contrast to the green fluorescence of the original. It was evaporated to dryness and the residue was dissolved in acetone and treated with powdered permanganate to oxidize any trace of acronycine. After excess permanganate had been destroyed with methanol the solution was filtered, concentrated to about 20 ml. and the hydrochloride of dihydroacronycine precipitated with concentrated hydrochloric acid. This was recrystallized from 10% hydrochloric acid giving orange plates, m.p. 124°C. (decomp.). *Dihydroacronycine*, liberated with ammonia, tended to darken on filtering at the pump. It crystallized from ethyl acetate-light petroleum as long yellow needles, m.p. 140-141.5°C.

Found: C, 74.3; H, 6.3; N, 4.1;  $CH_3O$ , 9.2%.

Calculated for  $C_{20}H_{21}O_3N$ : C, 74.3; H, 6.5; N, 4.3;  $CH_3O$ , 9.6% (one methoxyl).

(i) *norDihydroacronycine*.—Dihydroacronycine hydrochloride (2.3 g.) was heated in a test tube to its melting point and maintained at that temperature until effervescence ceased. The residue yielded yellow cubes of *nordihydroacronycine* from alcohol, m.p. 215-216°C.

Found: C, 73.7; H, 6.3; N, 4.5%.

Calculated for  $C_{18}H_{19}O_3N$ : C, 73.8; H, 6.2; N, 4.7%.

The same product was obtained by refluxing dihydroacronycine with hydrobromic acid in acetic acid for 2 hours or with hydrochloric acid in alcohol for a longer period.

#### (c) *Oxidation of Acronycine with Permanganate*

Acronycine (2 g.) was dissolved in acetone (100 ml.) and treated with finely powdered potassium permanganate (3.5 g.) with gentle warming. The oxidation was complete in an hour, when the precipitate was filtered and washed with a little acetone. It was then suspended in water (100 ml.), and treated with sulphur dioxide at 0°C., until the manganese dioxide dissolved. At the same time, a yellow solid was precipitated. This was filtered and redissolved in dilute caustic soda solution. A small amount of manganese dioxide was removed by filtration and the filtrate again acidified with sulphur dioxide at 0°C. to give a yellow solid which separated from aqueous methanol in pale yellow microcrystals, m.p. 215-217°C. Yield 0.9 g.

Found: C, 63.0; H, 5.1; N, 3.9%.

Calculated for  $C_{20}H_{19}O_7N$ : C, 62.3; H, 4.9; N, 3.6%.

When the precipitation of this acid with sulphur dioxide was carried out in the warm, or when the isolated acid was recrystallized from 10% hydrochloric acid, one mole of carbon dioxide was lost with the formation of a monobasic acid, m.p. 227°C. This acid has been termed *acronycinic acid*.

Found: C, 67.0; H, 5.5; N, 4.3;  $CH_3O$ , 9.4%.

$C_{19}H_{19}O_6N$  requires: C, 66.9; H, 5.6; N, 4.1;  $CH_3O$ , 9.1% (one methoxyl).

It was characterized by its methyl ester, prepared by the action of diazomethane in dioxan



solution or as follows: Acronycinic acid (1 g.) in methyl alcohol (20 ml.) was heated on a water-bath for 1 hour with concentrated sulphuric acid (1 ml.). On cooling, long yellow needles of the sulphate of the methyl ester separated. These were filtered, treated with ammonia, and recrystallized from methanol, giving the *methyl ester* in the form of pale, yellow needles, m.p. 132-133°C.

Found: C, 67.6; H, 6.3; N, 4.0;  $\text{CH}_3\text{O}$ , 17.4%.

Calculated for  $\text{C}_{20}\text{H}_{21}\text{O}_5\text{N}$ : C, 67.6; H, 5.9; N, 3.9;  $\text{CH}_3\text{O}$ , 17.5%.

#### (d) Pyrolysis of Acronycinic Acid

Acronycinic acid (2 g.) was placed in a small distilling flask with the side arm connected to a cooled receiver. The flask was heated in a bath to 240°C. until effervescence ceased. A small volatile fraction collected in the receiver and the residue in the flask set to a glass when cold.

(i) *The Volatile Fraction*.—This was dissolved in ether (10 ml.) and shaken in a separating funnel with dilute caustic soda (10 ml.) solution. The aqueous layer was separated, acidified with hydrochloric acid, and thoroughly extracted with ether. The ether extract was dried over sodium sulphate, filtered, and distilled. The residue was slowly sublimed under reduced pressure (20 mm.) when long white crystals collected in the cool arm of the flask. The combined yields from several experiments were recrystallized from light petroleum, giving  $\alpha$ -hydroxy-isobutyric acid, m.p. 76°C., undepressed on admixture with an authentic sample.

(ii) *The Non-Volatile Residue*.—The residue was treated with warm 10% caustic soda solution. The insoluble fraction was filtered and recrystallized from aqueous methanol (charcoal). The solution was decanted from the first dark crop of crystals and then yielded a crop of fine yellow crystals, which after further crystallization melted at 174-175°C.

Found: C, 70.9; H, 5.2; N, 5.6;  $\text{CH}_3\text{O}$ , 12.1%.

Calculated for  $\text{C}_{15}\text{H}_{13}\text{O}_3\text{N}$ : C, 70.6; H, 5.1; N, 5.5;  $\text{CH}_3\text{O}$ , 12.2% (one methoxyl).

The caustic soda solution above was acidified with carbon dioxide, yielding a brown precipitate. This was best purified by dissolving in 10% hydrochloric acid, heating with charcoal, and filtering. The clear filtrate was then diluted until a slight cloud developed. On standing, long straw coloured needles separated, m.p. 286-289°C. (decomp. uncorr.).

Found: C, 69.4; H, 4.7; N, 5.8%;  $\text{CH}_3\text{O}$ , nil.

Calculated for  $\text{C}_{14}\text{H}_{11}\text{O}_3\text{N}$ : C, 69.7; H, 4.6; N, 5.8%.

When this phenol, m.p. 286-289°C., was stirred with ether it slowly dissolved, and on recovery from this solution melted at 194-197°C. On analysis, however, the same formula was indicated.

Found: C, 70.1; H, 5.0; N, 5.8%.

Calculated for  $\text{C}_{14}\text{H}_{11}\text{O}_3\text{N}$ : C, 69.7; H, 4.6; N, 5.8%.

This material was dissolved in alkali and reprecipitated with carbon dioxide. It again melted at 286-289°C. (decomp. uncorr.).

The caustic soda insoluble substance,  $\text{C}_{15}\text{H}_{13}\text{O}_3\text{N}$ , (0.3 g.) dissolved in hot hydrobromic acid (10 ml.) was heated under reflux for 1 hour. On diluting, neutralizing with caustic soda, and acidifying with carbon dioxide, the phenol,  $\text{C}_{14}\text{H}_{11}\text{O}_3\text{N}$ , m.p. 286-289°C. (decomp. uncorr.), separated. The same phenol was obtained by heating acronycinic acid with hydrobromic acid under the same conditions as above.

The phenol,  $\text{C}_{14}\text{H}_{11}\text{O}_3\text{N}$ , (0.2 g.) dissolved in dioxan (10 ml.) was treated with a solution of diazomethane in dioxan, and allowed to stand overnight. Excess diazomethane was destroyed with acetic acid and the solvent removed under reduced pressure. The residue was crystallized from aqueous methanol and melted at 174-175°C. It was identical with the above "neutral" compound,  $\text{C}_{15}\text{H}_{13}\text{O}_3\text{N}$ , containing one methoxyl group.

Using dimethyl sulphate and concentrated alkali on the phenol,  $\text{C}_{14}\text{H}_{11}\text{O}_3\text{N}$ , in the usual way a dimethyl ether was obtained which after crystallization from chloroform-light petroleum had m.p. 161-163°C. This product exhibited a strong blue fluorescence in alcohol.

Found: N, 5.3%.

Calculated for  $\text{C}_{16}\text{H}_{15}\text{O}_3\text{N}$ : N, 5.2%.



(e) *Oxidation of norAcronycine*

*norAcronycine* was oxidized in acetone with potassium permanganate under the same conditions as for *acronycine*. No attempt was made to isolate the dibasic acid, but the monobasic acid, *noracronycinic acid*, was readily crystallized from aqueous methanol in small yellow prisms, m.p. 210-212°C.

Found: C, 66.3; H, 5.4; N, 4.3%.

Calculated for  $C_{18}H_{17}O_5N$ : C, 66.1; H, 5.2; N, 4.3%.

*norAcronycinic acid* heated under the conditions used above for *acronycinic acid* yielded  $\alpha$ -hydroxyisobutyric acid and the phenol,  $C_{14}H_{11}O_3N$ , m.p. 285-289°C. (decomp.).

(f) *Bromination of Acronycine*

*Acronycine* (2 g.) was dissolved in hot chloroform (20 ml.) and a bromine solution (6.5 ml.), prepared by dissolving bromine (16.5 g.) in chloroform (100 ml.) was added slowly with constant agitation. A bright red precipitate formed immediately in the hot solution. After standing overnight the crystals were filtered and sucked dry at the pump. The yield was 2.7 g. (90%) of crude hydrobromide, m.p. 135-137°C. (decomp.).

The hydrobromide was decomposed with dilute ammonia solution, the crude free base filtered, washed free of ammonia, and recrystallized from ethanol. This afforded golden-yellow needles of *monobromacronycine*, m.p. 195°C., solidifying, and remelting at 203°C.

Found: C, 59.8; H, 4.4; N, 3.4;  $CH_3O$ , 7.7%.

Calculated for  $C_{20}H_{18}O_3NBr$ : C, 60.0; H, 4.5; N, 3.5;  $CH_3O$ , 7.8% (one methoxyl).

The free base would not dissolve in hot 12% hydrochloric acid but dissolved readily in concentrated acid on warming. The hydrochloride crystallized in orange-red needles on cooling the latter solution. It melted with effervescence at 143°C. (decomp.) to yield a yellow compound, the latter remelting at 182°C. (decomp.).

(i) *Oxidation of Bromacronycine*.—*Bromacronycine* (1 g.) was dissolved in hot acetone (30 ml.) and powdered potassium permanganate (2 g.) added in small portions over half-an-hour. After a further refluxing for 10 minutes the manganese dioxide precipitate was filtered, washed with acetone, then mixed with ice water, and decomposed with sulphur dioxide. This afforded a yellow precipitate which was filtered, washed with water, and dissolved in dilute caustic soda solution on the water-bath. The resultant solution was filtered from a small amount of manganese dioxide and the acid again precipitated with sulphur dioxide. The precipitate was allowed to stand overnight and was then filtered, washed, and recrystallized from aqueous methanol, m.p. 227°C. It was identified as *acronycinic acid* by mixed melting point.

(ii) *Ozonolysis of Bromacronycine*.—*Bromacronycine* (1 g.) was dissolved in chloroform (50 ml.) and ozonized oxygen passed through the ice cold solution for 3 hours by which time a yellow suspension had separated. The chloroform was then evaporated and the residue refluxed with water for one hour. The residue was extracted with hot dilute caustic soda solution. No precipitate was formed when this solution was treated with carbon dioxide, but sulphur dioxide precipitated a yellow powder which was filtered and washed. After several crystallizations from alcohol small yellow crystals of a phenolic acid (I) were obtained with m.p. 210-210.5°C.

Found: C, 64.1; H, 4.4; N, 5.0%.

Calculated for  $C_{16}H_{13}O_5N$ : C, 64.2; H, 4.4; N, 4.7%.

(iii) *Dibromnoracronycine*.—*Bromacronycine* (4 g.) was dissolved in chloroform (40 ml.) and a bromine solution (10.5 ml.; 15.6 g. bromine in 100 ml. chloroform) added. After a few minutes the hot solution was concentrated on the water-bath until an orange solid began to separate. After cooling this was filtered and the filtrate further concentrated. The combined crystals (4.6 g.) were taken up in a mixture of chloroform (400 ml.) and ethanol (20 ml.) and filtered from a small residue. Deep yellow crystals separated from the filtrate. Further crops were obtained by concentration of this solution.

It appeared that the product was a mixture of a yellow and a red compound and a separation

was attempted. After the two forms had been separated by an extended series of crystallizations from acetone, in which the material was sparingly soluble (c. 0.5 g. per litre) it was found that the interconversion of the two forms could be effected by choosing appropriate conditions for crystallization. Analytical figures supported the conclusion that these were dimorphic forms of one product. For example, when the product was warmed with aqueous pyridine a red crystalline residue remained, and on cooling the filtrate orange-yellow crystals separated. Further treatment of the orange-yellow product with insufficient aqueous pyridine to dissolve it left more red residue. Recrystallization of the red material from chloroform-ethanol gave the yellow variety.

Found: Yellow—C, 48.8; H, 3.4; N, 3.2%;  $\text{CH}_3\text{O}$ , nil.

Red—C, 48.9; H, 3.1; N, 3.2%;  $\text{CH}_3\text{O}$ , nil.

Calculated for  $\text{C}_{19}\text{H}_{15}\text{O}_3\text{N Br}_2$ : C, 49.1; H, 3.3; N, 3.0%.

Neither form melted sharply even after many recrystallizations. Both decomposed at c. 209°C., when a moderate rate of heating was used, but the decomposition point was largely dependent upon the heating rate.

#### (g) Bromination of *nor*Acronycine and *Bromnor*acronycine

*nor*Acronycine (2 g.) was dissolved in chloroform (25 ml.) and a bromine solution (6.4 ml.: 15.6 g. bromine in 100 ml. chloroform) added during 5 minutes. An orange-red precipitate separated rapidly in the warm solution. The slimy hydrobromide precipitate was difficult to filter so it was found more convenient to add excess dilute ammonia to the reaction mixture after it had stood for an hour. Vigorous shaking decomposed the hydrobromide and gave a chloroform solution of the free base which was dried and concentrated. Alcohol was then added and the remaining chloroform removed on the steam-bath. *Bromnor*acronycine crystallized on cooling, in yellow plates, m.p. 174–175°C. Yield 1.85 g. (72%).

Found: C, 58.9; H, 4.3; N, 3.8%.

Calculated for  $\text{C}_{19}\text{H}_{15}\text{O}_3\text{N Br}$ : C, 59.1; H, 4.2; N, 3.6%.

*Bromnor*acronycine (1 g.) was dissolved in chloroform and treated with a large excess (3.1 g.) of bromine. The reaction mixture was allowed to stand overnight. The filtered product was identical with that from the bromination of bromacronycine showing the same phenomenon of dimorphism. It is clear therefore that the hydrogen bromide eliminated during the bromination of bromacronycine caused the elimination of the labile methyl group as methyl bromide with the formation of a *nore*compound.

#### (h) Alcoholic Potash on Acronycine

Acronycine (2 g.) was dissolved in absolute alcohol (50 ml.) containing potassium hydroxide (5 g.) and the solution refluxed on a steam bath for 6 hours. A yellow powder separated and was filtered. It was dissolved in a large volume of water and the solution treated with carbon dioxide. The resulting precipitate was filtered and crystallized from a large volume of alcohol from which it separated in bright yellow plates, m.p. 202°C. Yield 0.8 g.

Found: C, 74.7; H, 6.2; N, 4.4;  $\text{CH}_3\text{O}$ , 9.8%.

Calculated for  $\text{C}_{20}\text{H}_{19}\text{O}_3\text{N}$ : C, 74.8; H, 5.9; N, 4.4;  $\text{CH}_3\text{O}$ , 9.7% (for one methoxyl).

This phenol is isomeric with acronycine and is readily reconverted to acronycine. Unlike acronycine it was not affected by cold 10% hydrochloric acid but on heating it dissolved to a red solution which on cooling deposited acronycine hydrochloride, identified by conversion to *nor*-acronycine and by the formation of acronycine by basifying with ammonia and crystallizing the liberated base.

The mother liquor after removal of the potassium salt of the phenol, m.p. 202°C., was concentrated to remove alcohol and diluted with water. On passing carbon dioxide through the solution a yellow amorphous material, m.p. 146–150°C., was precipitated, which was soluble in aqueous alkali and failed to give an insoluble hydrochloride in acetone. It could not be obtained in crystalline form from any solvent.

Found: C, 67.6; H, 6.5; N, 4.0%.

Calculated for  $\text{C}_{20}\text{H}_{19}\text{O}_3\text{N}, 2\text{H}_2\text{O}$ : C, 67.5; H, 6.4; N, 3.9%.

This product was readily converted into acronycine by refluxing with aqueous alcohol or ethyl acetate or by heating with 10% hydrochloric acid.

Other experiments using more concentrated alkali (40%) on acronycine failed to yield any acetone.

### III. ACKNOWLEDGMENT

The work described in this paper was carried out in the Chemistry Department, University of Melbourne.

### IV. REFERENCES

- (1) LAHEY, F. N., and THOMAS, W. C.—*Aust. J. Sci. Res. A* **2**: 423 (1949).
- (2) CROW, W. D., and PRICE, J. R.—*Ibid.* **2**: 255 (1949).
- (3) BRIDGE, W., HEYES, R. G., and ROBERTSON, A.—*J. Chem. Soc.* **1937**: 279 (1937).
- (4) MCGOOKIN, A., REED, F. P., and ROBERTSON, A.—*Ibid.* **1937**: 748 (1937).
- (5) GEORGE, S. W., and ROBERTSON, A.—*Ibid.* **1937**: 1535 (1937).



# ALKALOIDS OF THE AUSTRALIAN RUTACEAE: *ACRONYCHIA BAUERI*

## III. THE STRUCTURE OF ACRONYCINE

By L. J. DRUMMOND and F. N. LAHEY\*

[Manuscript received August 5, 1949]

### Summary

Oxidative degradations of acronycine with nitric acid and ozone have been studied. Hot concentrated nitric acid converts *nora*acronycine and *nordi*hydroacronycine into 1-methyl-4-quinolone-3-carboxylic acid while nitric acid on acronycine under the same conditions yields first trinitroacronycine and finally 6-nitro-1-methyl-4-quinolone-3-carboxylic acid. Ozonolysis of acronycine gives a phenolic aldehyde,  $C_{16}H_{13}O_4N$ , which has been deduced to be 1- or 3-formyl-2-hydroxy-4-methoxy-10-methylacridone (V or VII). These data together with information supplied in Part II indicate acronycine to be either 4-methoxy-2',2',10-trimethyl- $\alpha$ -pyrano-(5',6'-1,2)acrid-5-one (IV) or 4-methoxy-2',2',10-trimethyl- $\alpha$ -pyrano-(5',6'-3,2)acrid-5-one (VI).

### I. INTRODUCTION

In Part II (1) it was shown that oxidation of acronycine with potassium permanganate established the presence of a *gem*dimethylpyran nucleus. More vigorous oxidation with nitric acid leads to a variety of products depending largely on experimental conditions. Concentrated nitric acid in alcohol (1:5) yields a mononitroacronycine,  $C_{20}H_{18}O_3N(NO_2)$ , the nitro group presumably entering the same position on the pyran ring as the bromine in bromacronycine(1) for on oxidation with potassium permanganate, nitroacronycine gives acronycinic acid.

With hot concentrated nitric acid acronycine is converted to an insoluble orange trinitroacronycine. By prolonged heating with excess concentrated nitric acid this trinitroacronycine gradually dissolves and gives rise to a cream crystalline acid,  $C_{11}H_8O_5N_2$ , m.p. 262-263°C. (uncorr.). *nor*Acronycine or *nordi*hydroacronycine with hot concentrated nitric acid gives an acid,  $C_{11}H_8O_3N$ , m.p. 296-297°C. (uncorr.). This acid proved to be identical with the acid of the same composition and melting point obtained by Price(2) by oxidation of the *Melicope* alkaloids. Price(2) produced evidence to show that this acid was 1-methyl-4-quinolone-3-carboxylic acid (I) and this has now been confirmed by synthesis. 4-Hydroxyquinoline-3-carboxylic acid was synthesized by the method of Gould and Jacobs(3) and was methylated to yield 1-methyl-4-quinolone-3-carboxylic acid identical with the product obtained by degradation of *nora*acronycine. The identity has been confirmed by comparison of the bases formed by decarboxylating both the naturally derived and synthetic acids. In addition we have synthesized 1,3-dimethyl-4-quinolone and this has been shown by Price (loc. cit.) to be identical with the product formed by direct reduction

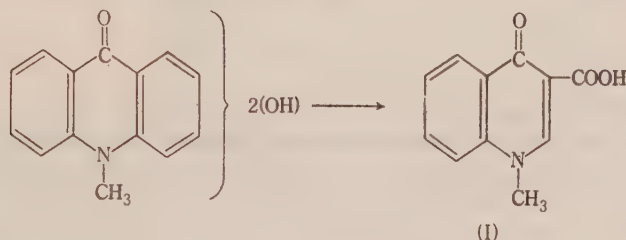
\*Present address: Chemistry Department, University of Queensland.



of the carboxyl group of 1-methyl-4-quinolone-3-carboxylic acid with zinc and hydrochloric acid.

The acid,  $C_{11}H_8O_5N_2$ , is clearly a nitro derivative of 1-methyl-4-quinolone-3-carboxylic acid and has been shown by synthesis to be 6-nitro-1-methyl-4-quinolone-3-carboxylic acid.

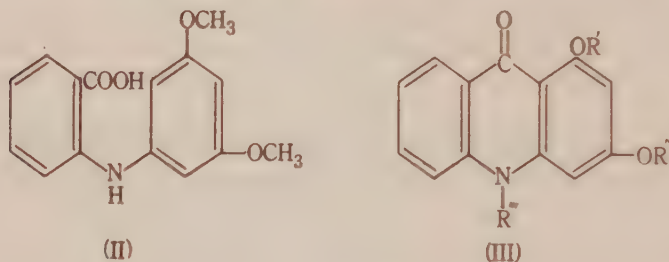
It has previously been shown(1) that the dimethylpyran ring present in acronycine can be removed to leave a compound,  $C_{14}H_{11}O_3N$ , containing two hydroxy groups, one of which is phenolic in character. The formation of 1-methyl-4-quinolone-3-carboxylic acid (I) by more vigorous oxidation with nitric acid can only be accounted for satisfactorily if this phenol contains an aromatic ring fused to the quinolone nucleus. Furthermore this ring must carry the two hydroxyl groups which are lost in the oxidation and the phenol must be a dihydroxy-10-methylacridone. Acronycine is thus another member of the new group of acridine alkaloids(4).



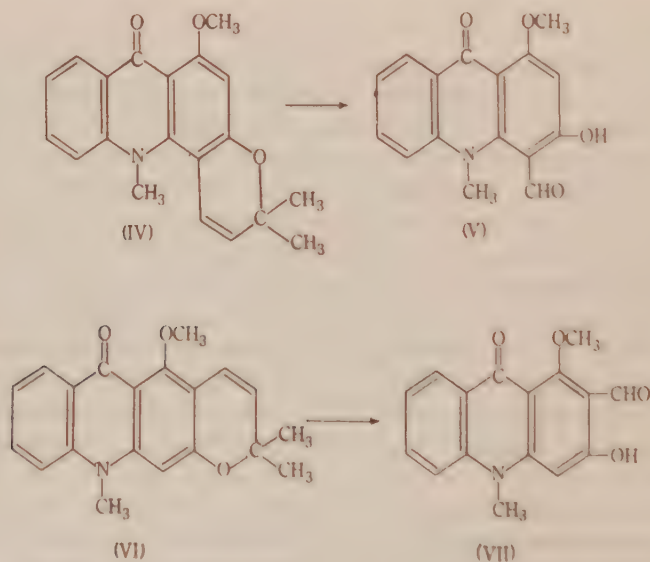
It was recorded in Part II (loc. cit.) that acronycine hydrochloride on being heated in the dry state is converted to *noracronycine*, the change simply involving the demethylation of a methoxyl group. The resulting hydroxyl group, however, shows no acidic properties and the change to *noracronycine* is accompanied by an enhancement of colour. This hydroxyl group is placed at position 4 on the acridone nucleus where hydrogen bonding with the carbonyl group would be expected, thereby affording an explanation of the reduced acidity of this hydroxyl group and the colour change on passing to the *nor*compound. The latter point will be discussed further in a separate paper when spectrographic evidence to support the position of this hydroxyl group will be offered. That hydrogen bonding does occur is supported by the observation of a small difference ( $4^{\circ}C.$ ) between the wet and dry melting points of *noracronycine*(5). Furthermore this allocation makes *noracronycine* analogous to the closely related *noralkaloids* of *Melicope fareana* studied by Crow and Price(4).

It seems reasonable to suppose that the hydroxy group showing no phenolic properties in the phenol,  $C_{14}H_{11}O_3N$ , is the same as that formed during the conversion of acronycine to *noracronycine*. The phenol fails to yield a quinone on oxidation under mild conditions, suggesting that the second hydroxyl group is not located at positions 1 or 3. Proof that this phenol is 2,4-dihydroxy-10-methylacridone was obtained by synthesis. Condensation of 3,5-dimethoxyaniline with *o*-chlorbenzoic acid gave 3',5'-dimethoxyphenylantranilic acid (II) which readily cyclized to 2,4-dimethoxyacridone (III,  $R', R'' = CH_3, R''' = H$ ). The potassium salt on methylation with dimethyl sulphate yielded 2,4-dimethoxy-10-methylacridone

(III, R', R'', R''' = CH<sub>3</sub>). The hydrochloride of this base on heating gave 4-hydroxy-2-methoxy-10-methylacridone (III, R' = H, R'', R''' = CH<sub>3</sub>) m.p. 174-175°C., identical with the non-phenolic base, C<sub>15</sub>H<sub>13</sub>O<sub>3</sub>N, obtained by the pyrolysis of acronycinic acid. Hydrobromic acid converted this base or the 2,4-dimethoxy-10-methylacridone to 2,4-dihydroxy-10-methylacridone (III, R', R'' = H, R''' = CH<sub>3</sub>) identical with the phenol, C<sub>14</sub>H<sub>11</sub>O<sub>3</sub>N. Ozonolysis of acronycine gave a

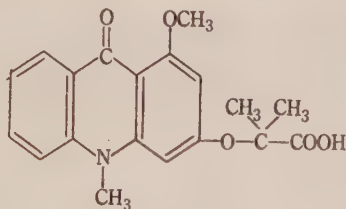


phenolic aldehyde, C<sub>16</sub>H<sub>13</sub>O<sub>4</sub>N, by rupture of the dimethylpyran ring. This aldehyde must therefore be 1- or 3-formyl-4-methoxy-2-hydroxy-10-methylacridone (V or VII) depending on whether acronycine is represented by either formula IV or VI.



As yet no definite evidence is available to enable us to decide between these two possibilities. Either structure would lead to the same formula for acronycinic acid (VIII) which on pyrolysis yields a mixture of 2,4-dihydroxy-10-methylacridone (III, R', R'' = H, R''' = CH<sub>3</sub>) and 4-hydroxy-2-methoxy-10-methylacridone (III, R' = H, R'', R''' = CH<sub>3</sub>). The formation of the latter rather than the isomeric 2-hydroxy-4-methoxy-10-methylacridone can be accounted for by assuming

either that after elimination of the  $\alpha$ -hydroxyisobutyric acid from the side chain the acridone ring opens and closes again yielding the more stable arrangement with the carbonyl group *ortho* to the hydroxyl instead of *ortho* to the methoxyl, or that the 2-hydroxy group of the phenol,  $C_{14}H_{11}O_3N$ , once formed is sufficiently acidic to react with more acronycinic acid causing demethylation of the 4-methoxyl group accompanied by the simultaneous formation of a methoxyl group at position 2 just as acronycine hydrochloride on being heated yields methyl chloride. The latter explanation appears to us to be the more likely.



(VIII)

The aldehyde (V or VII) obtained by ozonolysis of acronycine was methylated with dimethyl sulphate in acetone in the presence of potassium carbonate and yielded 1- or 3-formyl-2,4-dimethoxy-10-methylacridone which was oxidized to the corresponding carboxylic acid with potassium permanganate. This acid readily loses carbon dioxide to leave 2,4-dimethoxy-10-methylacridone, m.p. 161-163°C., identical with the synthetic product above (III, R', R'', R''' = CH<sub>3</sub>).

## II. EXPERIMENTAL

All melting points are corrected unless otherwise stated. Microanalyses are by Mr. N. L. Lottkowitz.

### (a) The Action of Nitric Acid on Acronycine

(i) *Dilute Nitric Acid*.—To a solution of acronycine (1 g.) in alcohol (20 ml.) was added concentrated nitric acid (4 ml.). The mixture was boiled gently for 1-2 minutes and poured into water (500 ml.). The bright yellow precipitate was filtered and crystallized from ethyl acetate yielding nitroacronycine, m.p. 222°C.

Found: C, 66.0; H, 4.9; N, 7.5; CH<sub>3</sub>O, 7.9%.

Calculated for C<sub>20</sub>H<sub>18</sub>O<sub>3</sub>N(NO<sub>2</sub>): C, 65.6; H, 4.9; N, 7.7; CH<sub>3</sub>O, 8.5% (for one methoxyl).

Nitroacronycine was oxidized with potassium permanganate in acetone solution and the product worked up as in the oxidation of acronycine(1). The same acid, acronycinic acid, m.p. 226°C. (decomp.) was obtained.

(ii) *Concentrated Nitric Acid*.—(a) Acronycine (1 g.) was gently heated with concentrated nitric acid (3 ml.) to about 50°C. A yellow solid quickly separated and when the reaction ceased the mixture was diluted with water and filtered. The bright yellow solid obtained was insoluble in all solvents tried. It was purified by refluxing in turn with ethanol, chloroform, and acetone. The remaining product (0.65 g.) melted at 290-291°C. (uncorr.).

Found: C, 53.1; H, 3.6; N, 12.2; CH<sub>3</sub>O, 6.9%.

Calculated for C<sub>20</sub>H<sub>16</sub>O<sub>3</sub>N(NO<sub>2</sub>)<sub>3</sub>: C, 52.6; H, 3.5; N, 12.2; CH<sub>3</sub>O, 6.8% (for one methoxyl).

(b) Acronycine (2 g.) was boiled with concentrated nitric acid (20 ml.) for about 2 hours, the solution being allowed to concentrate to a final volume of 5 ml. On addition of water a light



yellow solid was thrown down. It was soluble in bicarbonate and crystallized from glacial acetic acid. It was further purified by dissolving in alkali and treating with a little potassium permanganate solution. Sulphur dioxide then precipitated the acid again. On warming with 5% alkali it went into solution and on cooling the sodium salt separated. Acidification with sulphurous acid regenerated the acid which crystallized from acetic acid or aqueous dioxan in small cream plates, m.p. 262-263°C. (uncorr.).

Found: C, 53.4; H, 3.4; N, 11.5%.

Calculated for  $C_{11}H_8O_3N_2$ : C, 53.3; H, 3.2; N, 11.3%.

This acid was shown to be 6-nitro-1-methyl-4-quinolone-3-carboxylic acid by comparison with a synthetic sample (see below).

Decarboxylation of this acid was carried out in dibutyl phthalate with a little copper powder. The cooled solution was diluted with chloroform, extracted three times with 10% hydrochloric acid and the aqueous layer made alkaline with potash and continuously extracted with chloroform. After removing the solvent, the residue was crystallized from methanol, and then acetic acid, yielding yellow needles, m.p. 238-239°C.

Found: N, 13.8%.

Calculated for  $C_{10}H_8O_3N_2$ : N, 13.7%.

This product gave no depression of melting point when mixed with a synthetic specimen of 6-nitro-1-methyl-4-quinolone.

#### (b) The Action of Nitric Acid on *nordihydroacronycine* and *norAcronycine*

Concentrated nitric acid (4.5 ml.) was added to *nordihydroacronycine* (0.8 g.) and the mixture gently warmed. After the initial vigorous reaction had ceased, the solution was taken almost to dryness, and poured into ice water to precipitate a yellow acid which was purified by treatment with a little potassium permanganate in alkaline solution. On acidification with sulphur dioxide a light yellow acid was thrown down. On crystallization from glacial acetic acid, it was obtained as colourless needles which melted at 296-297°C. (uncorr.) with softening at c. 280°C. This acid was identified as 1-methyl-4-quinolone-3-carboxylic acid (I) by comparison with a synthetic specimen.

Oxidation of *noracronycine* under the same conditions as for *nordihydroacronycine* also yielded 1-methyl-4-quinolone-3-carboxylic acid.

Decarboxylation of this acid carried out as above for 6-nitro-1-methyl-4-quinolone-3-carboxylic acid yielded 1-methyl-4-quinolone, m.p. 152-153°C., identified by mixed melting point with a synthetic specimen. In addition to the comparison of the bases from synthetic and natural sources the following salts were compared by mixing melting point determinations and in every case no depression was observed: hydrochloride, m.p. 196-197°C., picrate, m.p. 231-232°C., platinichloride, m.p. 219.5-220.5°C. (decomp.).

#### (c) 1-Methyl-4-Quinolone-3-Carboxylic Acid

The synthesis of 3-carboxy-4-hydroxyquinoline was carried out according to the procedure of Gould and Jacobs(3). Methylation: (i) The ester (1.0 g.) suspended in 50% caustic soda solution (30 ml.) was shaken with dimethylsulphate (15 ml.) added in small portions. The white solid which separated was filtered and dissolved in water. This solution was extracted with chloroform and then acidified to Congo red with dilute hydrochloric acid. The precipitated 1-methyl-4-quinolone-3-carboxylic acid (I) was filtered, washed with water, and crystallized from 60% acetic acid. It melted at 296-297°C. (uncorr.) with softening at c. 280°C.

Found: N, 7.0%;  $CH_2O$ , nil.

Calculated for  $C_{11}H_8O_3N$ : N, 6.9%.

(ii) To the ester (4.5 g.) in methanol (10 ml.) containing sodium (0.6 g.) was added methyl iodide (10 ml.). The tube was sealed and held at 100°C. for 26 hours. On working up in the usual way 1-methyl-4-quinolone-3-carboxylic acid, m.p. 296-297°C. (uncorr.) was obtained.

#### (d) 1-Methyl-4-Quinolone

The above synthetic acid was decarboxylated by the procedure previously described, to yield 1-methyl-4-quinolone, m.p. 152-153°C., after crystallization from benzene.



Found: C, 75.8; H, 5.8; N, 8.8%.

Calculated for  $C_{10}H_9ON$ : C, 75.5; H, 5.7; N, 8.8%.

Alternatively the synthetic 4-hydroxyquinoline-3-carboxylic acid was decarboxylated to 4-hydroxyquinoline after Gould and Jacobs (loc. cit.) and this was N-methylated as follows: 4-hydroxyquinoline (0.6 g.) was added to methanol (1.5 ml.) in which was dissolved sodium (0.1 g.). After adding methyl iodide (1 ml.) the tube was sealed and heated at 100°C. for 20 hours. The tube was opened and the contents poured into water. Chloroform extraction of this gave a product which was taken up in benzene. The first crop of dark material was filtered and discarded. Careful addition of light petroleum to the filtrate gave 1-methyl-4-quinolone as white needles, m.p. 152-153°C.

(e) 6-Nitro-1-Methyl-4-Quinolone-3-Carboxylic Acid

Following the procedure of Riegel *et al.* (6) 6-nitro-4-hydroxyquinoline-3-carboxylic acid was prepared. This acid (3 g.) was placed in a pyrex tube with methanol (4 ml.) in which was dissolved sodium (0.2 g.) and methyl iodide (3 ml.). The tube was sealed and heated at 100°C. for 19 hours. After opening the tube the contents were washed out with 5% alkali (20 ml.) and refluxed for 15 minutes. The solution was filtered hot and on cooling the sodium salt of 6-nitro-1-methyl-4-quinolone-3-carboxylic acid settled out. It was redissolved in hot water, treated with charcoal, filtered, and acidified to precipitate a pale yellow acid. After crystallization from aqueous dioxan or acetic acid it melted at 262-264°C. (uncorr.).

Found: N, 11.3%.

Calculated for  $C_{11}H_8O_5N_2$ : N, 11.3%.

(f) 6-Nitro-1-Methyl-4-Quinolone

Decarboxylation of this acid in dibutyl phthalate gave 6-nitro-1-methyl-4-quinolone, m.p. 238-239°C.

Found: N, 13.8%.

Calculated for  $C_{10}H_8O_3N_2$ : N, 13.7%.

(g) 1,3-Dimethyl-4-Quinolone

By the method of Steck, Hallock, and Holland (7) 3-methyl-4-hydroxyquinoline, m.p. 228-230°C., was prepared. This base (1.4 g.) in methanol (4 ml.) containing sodium (0.2 g.) was sealed in a pyrex tube with methyl iodide (3 ml.) and held at 100°C. for 24 hours. The product isolated in the usual way was crystallized from benzene-light petroleum, then sublimed under reduced pressure. Further crystallization from benzene-light petroleum gave colourless needles, m.p. 152-153°C.

Found: C, 76.1; H, 6.3; N, 8.0%.

Calculated for  $C_{11}H_{11}ON$ : C, 76.3; H, 6.4; N, 8.1%.

The picrate of 1,3-dimethyl-4-quinolone formed yellow needles, m.p. 189-190°C.

(h) 4-Hydroxy-2-Methoxy-10-Methylacridone

(i) 3,5-Dimethoxyaniline.—5-Nitroresorcinol dimethylether (8) (8.4 g.) was dissolved in ethyl alcohol (150 ml.) and shaken with Raney nickel in an atmosphere of hydrogen at atmospheric pressure. When absorption of gas had ceased, the catalyst was filtered off, and the filtrate concentrated. The residue was poured into 10% hydrochloric acid and the solution extracted with ether. After making alkaline with caustic soda the solution was again extracted with ether. The ether extract was dried over sodium sulphate and the solvent removed. The residual oil was distilled under reduced pressure yielding a colourless oil which on cooling solidified, m.p. 66-68°C.

(ii) 3',5'-Dimethoxyphenylanthranilic Acid (II).—3,5-Dimethoxyaniline (2 g.) and *o*-chlorbenzoic acid (2 g.) in amyl alcohol (20 ml.) were refluxed in the presence of potassium carbonate and a little copper catalyst in an atmosphere of carbon dioxide. The solvent was then removed by steam distillation and the alkaline solution clarified with charcoal. Acidification precipitated 3',5'-dimethoxyphenylanthranilic acid which crystallized from aqueous ethanol in fine greenish-yellow needles, m.p. 147°C.

Found: C, 65.7; H, 5.8; N, 5.1;  $\text{CH}_3\text{O}$ , 22.7%.

Calculated for  $\text{C}_{18}\text{H}_{15}\text{O}_4\text{N}$ : C, 65.9; H, 5.5; N, 5.1;  $\text{CH}_3\text{O}$ , 22.7%.

(iii) 2,4-Dimethoxyacridone (III,  $R'$ ,  $R'' = \text{CH}_3$ ,  $R''' = \text{H}$ ).—3',5'-Dimethoxyphenyl-anthranilic acid (0.5 g.) was refluxed with phosphorus oxychloride (5 ml.) for half-an-hour. The red solution was poured into cold water and the solution filtered. Neutralization with ammonia precipitated 2,4-dimethoxyacridone which crystallized from aqueous ethanol in fine yellow needles, m.p. 162°C.

Found: N, 5.5;  $\text{CH}_3\text{O}$ , 24.2%.

Calculated for  $\text{C}_{18}\text{H}_{15}\text{O}_4\text{N}$ : N, 5.5;  $\text{CH}_3\text{O}$ , 24.3%.

(iv) 2,4-Dimethoxy-10-Methylacridone (III,  $R'$ ,  $R''$ ,  $R''' = \text{CH}_3$ ).—2,4-Dimethoxyacridone (0.3 g.) was dissolved in dry methanol (10 ml.) and added to a solution of methanolic potash (0.25 g. KOH). The alcohol was removed on the steam bath and the yellow solid was treated with an excess of dimethyl sulphate. After the initial vigorous reaction the whole was held at 100°C. for half-an-hour. Excess dimethyl sulphate was destroyed with caustic soda and the precipitated 2,4-dimethoxy-10-methylacridone filtered from the cold solution. It crystallized from aqueous ethanol or chloroform-light petroleum in fine colourless needles, m.p. 161-163°C.

Found: N, 5.3;  $\text{CH}_3\text{O}$ , 22.7;  $\text{CH}_3\text{N}$ , 8.3%.

Calculated for  $\text{C}_{19}\text{H}_{17}\text{O}_4\text{N}$ : N, 5.2;  $\text{CH}_3\text{O}$ , 23.0;  $\text{CH}_3\text{N}$ , 10.8% (for two methoxyl and one N-methyl groups).

(v) 4-Hydroxy-2-Methoxy-10-Methylacridone (III,  $R' = \text{H}$ ,  $R''$ ,  $R''' = \text{CH}_3$ ).—A sample of 2,4-dimethoxy-10-methylacridone treated with 10% hydrochloric acid gave a hydrochloride which crystallized in yellow needles. It was filtered and dried over sulphuric acid. It melted at 133°C. with decomposition. A sample of the hydrochloride was treated in a test tube in an oil bath at 140°C. Effervescence occurred and the melt gave yellow needles on cooling. After recrystallization from aqueous ethyl alcohol the 4-hydroxy-2-methoxy-10-methylacridone melted at 174-175°C.

Found: C, 70.5; H, 5.6; N, 5.7%.

Calculated for  $\text{C}_{19}\text{H}_{17}\text{O}_4\text{N}$ : C, 70.6; H, 5.1; N, 5.5%.

No depression of the melting point was observed when this product was mixed with the non-phenolic base (of the same melting point) formed by the thermal decomposition of acronycine acid (Part II loc. cit.).

(i) 2,4-Dihydroxy-10-Methylacridone (III,  $R'$ ,  $R'' = \text{H}$ ,  $R''' = \text{CH}_3$ )

2,4-Dimethoxy-10-methylacridone (0.5 g.) was refluxed for 1 hour with hydrobromic acid (7 ml.) by which time a greenish precipitate had settled out. The whole was poured into water and warmed on a steam bath to coagulate the precipitate, which was then filtered, dissolved in alkali, and reprecipitated with carbon dioxide. From aqueous ethanol the 2,4-dihydroxy-10-methylacridone was obtained as a fine yellow powder, m.p. 286-289°C. (decomp. uncorr.) with shrinking beginning at 270°C.

Found: N, 5.8%;  $\text{CH}_3\text{O}$ , nil.

Calculated for  $\text{C}_{19}\text{H}_{17}\text{O}_4\text{N}$ : N, 5.8%.

By mixed melting point determinations it was shown to be identical with the phenol of the same melting point obtained by degradation of acronycine acid (Part II loc. cit.).

(j) Ozonolysis of Acronycine

Acronycine (4 g.) was dissolved in a mixture of acetic acid (50 ml.) and ethyl acetate (2 ml.), cooled in ice, and treated with a stream of ozone. The completion of the reaction was indicated by a change in colour of the solution from red to yellow accompanied by the precipitation of a fine yellow powder. This was filtered, washed, boiled with water for 1 hour, and the filtered product dissolved in alkali. Carbon dioxide precipitated a small amount of a yellow powder which could not be satisfactorily crystallized. It was not further examined. Sulphur dioxide then precipitated the main product 1- or 3-formyl-2-hydroxy-4-methoxy-10-methylacridone (V or VII) as a pale yellow solid which crystallized from alcohol in needles, m.p. 235°C. With ferric chloride it gave a bright red colouration.

Found: C, 67.8; H, 4.8; N, 5.1;  $\text{CH}_3\text{O}$ , 11.0%.

Calculated for  $\text{C}_{16}\text{H}_{15}\text{O}_4\text{N}$ : C, 67.8; H, 4.6; N, 5.1;  $\text{CH}_3\text{O}$ , 11.0%.

This phenolic aldehyde (0.4 g.) dissolved in acetone (50 ml.) was refluxed over potassium carbonate, and dimethyl sulphate (6 ml.) added in small portions. After 2 hours the solution was filtered and concentrated on the water-bath. When a yellow solid commenced to settle in the hot, water was added and then 10% caustic soda solution. The precipitated 1- or 3-formyl-2,4-dimethoxy-10-methylacridone was filtered and crystallized from aqueous methanol (charcoal), m.p. 217-218°C. It gave no colour with ferric chloride.

Found: C, 68.4; H, 5.2; N, 4.7%.

Calculated for  $\text{C}_{17}\text{H}_{15}\text{O}_4\text{N}$ : C, 68.7; H, 5.1; N, 4.7%.

The methylated aldehyde (0.4 g.) suspended in sodium carbonate solution was oxidized to the corresponding carboxylic acid with potassium permanganate (0.5 g.). The acid isolated from this solution in the usual way melted at 195-196°C. It was decarboxylated by heating in dibutyl phthalate at 150°C. until the evolution of carbon dioxide ceased. Ether precipitated 2,4-dimethoxy-10-methylacridone in almost colourless cubes, m.p. 161-163°C. In dilute alcohol solution it exhibited a brilliant purple fluorescence.

Found: N, 5.1%.

Calculated for  $\text{C}_{16}\text{H}_{15}\text{O}_3\text{N}$ : N, 5.2%.

A mixed melting point of this base and the synthetic 2,4-dimethoxy-10-methylacridone established the identity of the two. In addition the picrates, m.p. 203-205°C. (decomp.) and hydrochlorides, m.p. 133°C. (decomp.) of the bases from synthetic and natural sources were found to be identical.

### III. ACKNOWLEDGMENT

The work described in this paper was carried out in the Chemistry Department, University of Melbourne.

### IV. REFERENCES

- (1) BROWN, R. D., DRUMMOND, L. J., LAHEY, F. N., and THOMAS, W. C.—*Aust. J. Sci. Res.* A 2: p. 622 (1949).
- (2) PRICE, J. R.—*Ibid.* 2: 272 (1949).
- (3) GOULD, R. G., and JACOBS, W. A.—*J. Amer. Chem. Soc.* 61: 2890 (1939).
- (4) CROW, W. D., and PRICE, J. R.—*Aust. J. Sci. Res.* A 2: 282 (1949).
- (5) BAKER, W.—*J. Chem. Soc.* 1934: 1685 (1934).
- (6) RIEGEL, B., *et al.*—*J. Amer. Chem. Soc.* 68: 1264 (1946).
- (7) STECK, E. A., HALLOCK, L. L., and HOLLAND, A. J.—*Ibid.* 68: 129 (1946).
- (8) VERMEULEN, H.—*Rec. trav. Chim.* 25: 27 (1906).





# INDEX

	PAGE
<i>Acacia acuminata</i> , Occurrence of $\beta$ -Carotene and "Phytofluene" in the Wood of ...	132
Acetylacetone, Metal Complexes of ...	92
<i>Acronychia Baueri</i> , Isolation of the Alkaloids of ...	423
<i>Acronychia Baueri</i> , Alkaloids of Australian Rutaceae ...	622, 630
Acronycine, Some Reactions of Alkaloid ...	622
Acronycine, Structure of ...	630
Alkali Lignin ...	600
Alkaloids, A Degradation Product of ...	272
Alkaloids of the Australian Rutaceae : <i>Acronychia Baueri</i> ...	423, 622, 630
Alkaloids of the Australian Rutaceae : <i>Evodia xanthoxyloides</i> ...	429
Alkaloids of the Australian Rutaceae : <i>Melicope fareana</i> ...	249-82
Alkaloids, Isolation of ...	423
Alkaloids, Structure of ...	282
Anet, E. F. L. J., Hughes, G. K., and Ritchie, E.— A Synthesis of Hygrine and Cuscohygrine ...	616
Anet, F. A. L., Blanks, F. R., and Hughes, G. K.— The Chemical Constituents of Australian <i>Flindersia</i> Species. I. Collinin - 7 - Geranoxy - 8 - Methoxycoumarin ...	127
Anet, F. A. L., Hughes, G. K., and Ritchie, E.— The Chemical Constituents of Australian <i>Flindersia</i> Species. II. Braylin and Brayleyanin...	608
Aniline in the Presence of Mercuric Oxide, the Action of Iodine on ...	111
Aromatic Compounds, Iodination of ...	111, 246, 595
Aromatic Ethers, Halogenation of ...	595

	PAGE
Bailey, V. A., and Roberts, J. A.— Graphical Study of the Dispersion of Electro-Magneto-Ionic Waves ...	307
Batchelor, G. K.— Diffusion in a Field of Homogeneous Turbulence. I. Eulerian Analysis ...	437
Binary Systems of Ketones and Hydrocarbons, Dielectric Properties of ...	405
Blanks, F. R.— <i>See</i> Anet, F. A. L., and Hughes, G. K.	
Bolton, J. G., and Stanley, G. J.— The Position and Probable Identification of the Source of Galactic Radio-Frequency Radiation Taurus-A ...	139
Bond Localization Energy ...	564
Bond Order, Definition, Methods of Computation, and Relation to ...	564
Bosworth, R. C. L.— The Second Viscosity Coefficient in Rheological Systems ...	394
Bower, J. C.— <i>See</i> Martin, L.H., <i>et al.</i>	
Brayleyanin ...	608
Braylin ...	608
Brown, R. D.— Bond Localization Energy. I. Definition, Methods of Computation, and Relation to Bond Order ...	564
Brown, R. D., Drummond, L. J., Lahey, F. N., and Thomas, W. C.— Alkaloids of the Australian Rutaceae : <i>Acronychia Baueri</i> . II. Some Reactions of the Alkaloid Acronycine ...	622
Bursts of Solar Noise ...	322
Bursts of Solar Radiation at Metre Wavelengths ...	214

	PAGE		PAGE
$\beta$ -Carotene and "Phytofluene" in the Wood of <i>Acacia acuminata</i> ... ..	132	Eclipse Observations ... ..	524
Christiansen, W. N., Yabsley, D. E., and Mills, B. Y.— Measurements of Solar Radiation at a Wavelength of 50 Centimetres during the Eclipse of November 1, 1948 ...	506	Eclipse of November 1, 1948, Solar Radiation during ...	506
Cloud Drop Distributions, Theoretical Aspects of ... ..	376	Elasticity in Curvilinear Coordinates, Equations of ...	483
Collinin-7-Geranoxy-8-Methoxycoumarin from the Australian <i>Flindersia</i> Species ... ..	127	Electromagnetic Radiation in an Ionized Medium, The Wave Equations for ... ..	169
Cosmic Ray Measurements between Australia and Japan ...	493	Electro-Magneto-Ionic Waves, Graphical Study of the Dispersion of ... ..	307
Cosmic Ray Showers, The Structure of ... ..	184	Emulsions, Contraction of Nuclear Energy, Bond Localization ...	564
Coumarin ... ..	127	Equations of Elasticity ...	483
Crow, W. D.— Alkaloids of the Australian Rutaceae: <i>Melicope fareana</i> . III. Oxidative Demethylation of Melicopine ... ..	264	Ethers, Halogenation of Aromatic ...	595
Crow, W. D., and Price, J. R.— Alkaloids of the Australian Rutaceae: <i>Melicope fareana</i> . II. Preliminary Examination of Melicopine, Melicopidine, and Melicopine ... ..	255	Ethers, Iodination of Aromatic ...	246
V. The Structure of the Alkaloids ... ..	282	<i>Eucalyptus regnans</i> , Lignin of ...	117, 600
Cuscohygrine, Synthesis of ...	616	Eulerian Analysis ... ..	437
Cylinder, Fully Developed Turbulent Wake of a ... ..	451	<i>Evodia xanthoxyloides</i> ... ..	429
Demethylation of Melicopine, Oxidative ... ..	264	Evoxanthine ... ..	429
Dielectric Properties of Binary Systems of Ketones and Hydrocarbons ... ..	405	Flexure of an Incomplete Tore ...	469
Diffusion in a Field of Homogeneous Turbulence ... ..	437	Flight Record, A Remarkable Thunderstorm ... ..	550
Drummond, L. J.—See Brown, R. D., et al.		<i>Flindersia</i> Species (Australian), the Chemical Constituents of ...	127, 608
Drummond, L. J., and Lahey, F. N.— Alkaloids of the Australian Rutaceae: <i>Acronychia Baueri</i> . III. The Structure of Acronychine ... ..	630	Freiberger, W.— On the Solution of the Equilibrium Equations of Elasticity in General Curvilinear Coordinates ... ..	483
Dunbar, D. N. F.—See Martin, L.H., et al.		The Uniform Torsion of an Incomplete Tore ... ..	354
		Freiberger, W., and Smith, R. C. T.— The Uniform Flexure of an Incomplete Tore ... ..	469
		Galactic Radio-Frequency Radiation Taurus-A, Position and Probable Identification of the Source of ... ..	139
		Geophysical Data, Practical Determination of Lunar and Luni-Solar Daily Variations in ...	1
		Giovanelli, R. G.— Hall Currents and the Ejection of Prominences by Sunspots ...	39
		Hall Currents and the Ejection of Prominences by Sunspots ...	39
		Halogenation of Aromatic Ethers ...	595
		He <sup>4</sup> and Li <sup>7</sup> ( $p, \alpha$ ) Reaction, $\alpha$ -Particles from ... ..	25

PAGE	PAGE
Hindman, J. V.—See Piddington, J. H.	
Hirst, F.—See Martin, L. H., <i>et al.</i>	
Histidine ... ..	579
Hovea Species, Isolation of <i>d</i> -Sparteine from ... ..	427
Hughes, G. K.—See Anet, E. F. L. J., and Ritchie, E.	
Hughes, G. K.—See Anet, F. A. L., and Blanks, F. R.	
Hughes, G. K.—See Anet, F. A. L., and Ritchie, E.	
Hughes, G. K., and Neill, K. G.—Alkaloids of the Australian Rutaceae: <i>Evodia xanthoxyloides</i> . I. Evoxanthine ... ..	429
Hydrocarbons, Dielectric Properties of Binary Systems of	405
Hydrogen Peroxide, Halogenation of Aromatic Ethers in the Presence of ... ..	595
8-Hydroxyquinoline, Metal Complexes of ... ..	92
8-Hydroxyquinoline 5-Sulphonic Acid, Metal Derivatives of ... ..	579
Hygrine and Cuscohygrine, Synthesis of ... ..	616
Ice, the Artificial Stimulation of Precipitation by Dry ... ..	232
Iodine on Aniline in the Presence of Mercuric Oxide, the Action of ... ..	111
Iodination of Aromatic Compounds ... ..	111, 246, 595
Jaeger, J. C., and Westfold, K. C.—Transients in an Ionized Medium with Applications to Bursts of Solar Noise ... ..	322
Jurd, L.—The Iodination of Aromatic Compounds. I. The Action of Iodine on Aniline in the Presence of Mercuric Oxide ... ..	111
II. The Iodination of Aromatic Ethers and of 2-Naphthol ... ..	246
III. Halogenation of Aromatic Ethers in the Presence of Hydrogen Peroxide ... ..	595
Ketones, Dielectric Properties of Binary Systems of ... ..	405
Kraus, E. B., and Smith, Betty.—Theoretical Aspects of Cloud Drop Distributions ... ..	376
Lahey, F. N.—See Brown, R. D., <i>et al.</i>	
Lahey, F. N.—See Drummond, L. J.	
Lahey, F. N., and Thomas, W. C.—Alkaloids of the Australian Rutaceae: <i>Acronychia Baueri</i> . I. The Isolation of the Alkaloids ... ..	423
Law, P. G., McKenzie, C. D., and Rathgeber, H. D.—Cosmic Ray Measurements between Australia and Japan	493
Lehany, F. J., and Yabsley, D. E.—Solar Radiation at 1200 Mc/s., 600 Mc/s., and 200 Mc/s. ... ..	48
Li <sup>7</sup> ( <i>p</i> , $\alpha$ ) He <sup>4</sup> Reaction, $\alpha$ -Particles from ... ..	25
Light from Optical Systems, The Distribution of ... ..	335
Lignin from Sulphate Pulp	117
Lignin, Alkali ... ..	600
Lignin of <i>E. regnans</i> ... ..	600
Lunar and Luni-Solar Variations in Geophysical Data ... ..	1
McKenzie, C. D.—See Law, P. G., and Rathgeber, H. D.	
Magnetic Field, The Wave Equations for Radiation in an Ionized Medium in a ... ..	169
Maley, L. E., and Mellor, D. P.—The Relative Stability of Internal Metal Complexes. I. Complexes of 8-Hydroxyquinoline, Salicylaldehyde, and Acetylacetone ... ..	92
II. Metal Derivatives of 8-Hydroxyquinoline 5-Sulphonic Acid, and a Series of Monocarboxylic Mono- $\alpha$ -Amino Acids including Histidine ... ..	579
Martin, L. H., Bower, J. C., Dunbar, D. N. F., and Hirst, F.—Angular Distribution of $\alpha$ -Particles from the Li <sup>7</sup> ( <i>p</i> , $\alpha$ ) He <sup>4</sup> Reaction ... ..	25
Martin, S. L.—The Contraction of Nuclear Emulsions ... ..	389



	PAGE		PAGE
Meakins, R. J.—		Noise, Bursts of Solar ...	322
The Dielectric Properties of		Noise-like Character of Solar	
Binary Systems of Ketones and		Radiation at Metre Wave-	
Hydrocarbons ...	405	lengths ...	228
<i>Melicope fareana</i> , Alkaloids of		Nuclear Emulsions, Contraction	
the Australian Rutaceae:	249-82	of ...	389
Melicopicine, Preliminary Ex-		Oxidative Demethylation of	
amination of ...	255	Melicopicine ...	264
Melicopicine, Oxidative Demethy-		$\alpha$ -Particles from the $\text{Li}^7(p, \alpha)\text{He}^4$	
lation of ...	264	Reaction, Distribution of ...	25
Melicopidine, Preliminary Ex-		Pawsey, J. L., and Yabsley,	
amination of ...	255	D. E.—	
Melicopine, Examination of ...	255	Solar Radio-Frequency Radia-	
Mellor, D. P.—See Maley, L. E.		tion of Thermal Origin ...	198
Mercuric Oxide, Iodine on Aniline		Payne-Scott, Ruby.—	
in the Presence of ...	111	Bursts of Solar Radiation at	
Mewether, J. W. T.—		Metre Wavelengths ...	214
Studies on the Lignin of <i>Euca-</i>		The Noise-like Character of	
lyptus regnans. III. The Iso-		Solar Radiation at Metre Wave-	
lation of an Alkali Lignin from		lengths ...	228
Sulphate Pulping ...	117	“Phytofluene” in the Wood of	
IV. Alkali Lignin ...	600	<i>Acacia acuminata</i> ...	132
Metal Complexes, Stability Con-		Piddington, J. H., and Hindman,	
stants of Internal ...	92, 579	J. V.—	
Metal Derivatives of 8-Hydroxy-		Solar Radiation at a Wave-	
quinoline 5-Sulphonic Acid		length of 10 Centimetres in-	
and a Series of Monocarboxylic		cluding Eclipse Observations...	524
Mono- $\alpha$ -Amino Acids ...	379	Piddington, J. H., and Minnett,	
1-Methyl-4-Quinolone-3-Carboxy-		H. C.—	
lic Acid, Some Reactions of ...	272	Microwave Thermal Radiation	
Metre Wavelengths, Bursts of		from the Moon ...	63
Solar Radiation at ...	214	Solar Radiation of Wavelength	
Metre Wavelengths, Noise-like		1.25 Centimetres ...	539
Character of Solar Radiation at	228	Precipitation by Dry Ice, The	
Microwave Thermal Radiation		Artificial Stimulation of ...	232
from the Moon ...	63	Prescott, J. R., and Mohr, C. B. O.	
Mills, B. Y.—See Christiansen,		The Structure of Cosmic Ray	
W. N., and Yabsley, D. E.		Air Showers ...	184
Minnett, H. C.—See Piddington,		Price, J. R.—	
J. H.		Alkaloids of the Australian	
Mohr, C. B. O.—See Prescott,		Rutaceae: <i>Melicope fareana</i> .	
J. R.		I. Isolation of the Constituent	
Monocarboxylic Mono- $\alpha$ -Amino		Alkaloids ...	249
Acids, Metal Derivatives of ...	579	IV. Some Reactions of 1-	
Moon, Microwave Thermal Radia-		Methyl-4-Quinolone-3-Carboxy-	
tion from ...	63	lic Acid, a Degradation Product	
Morrison, J. H., and Neill, K. G.—		of the Alkaloids ...	272
The Isolation of <i>d</i> -Sparteine		Price, J. R.—See Crow, W. D.	
from <i>Hovea</i> Species ...	427	Radiation at a Wavelength of 10	
2-Naphthol, Iodination of ...	246	Centimetres including Eclipse	
Neill, K. G.—See Hughes, G. K.		Observations, Solar ...	524
Neill, K. G.—See Morrison, J. H.			



	PAGE		F
Radiation at Wavelength of 50 Centimetres during Eclipse of November 1, 1948, Solar ...	506	<i>d</i> -Sparteine from <i>Hovea</i> Species, Isolation of ...	427
Radiation at 1200, 600, and 200 Mc/s., Solar ...	48	Squires, P., and Smith, E. J.—The Artificial Stimulation of Precipitation by Dry Ice ...	232
Radiation at Metre Wavelengths, Noise-like Character of Solar Radiation at Metre Wavelengths, Bursts of Solar ...	228	Stanley, G. J.—See Bolton, J. G.	
Radiation from the Moon, Microwave Thermal ...	63	Stanley, P., and Trikojus, V. M.—The Occurrence of $\beta$ -Carotene and "Phytofluene" in the Wood of <i>Acacia acuminata</i> ...	132
Radiation of Thermal Origin, Solar Radio-Frequency ...	198	Steel, W. H.—The Distribution of Light from Optical Systems ...	335
Radiation of Wavelength 1.25 Centimetres, Solar ...	539	Sulphate Pulping, Lignin from Sunspots, Hall Currents and the Ejection of Prominences by ...	117 39
Radiation in an Ionized Medium in a Magnetic Field, Wave Equations for Electromagnetic Radok, U.—A Remarkable Thunderstorm Flight Record ...	169 550	Taurus-A, Source of Galactic Radio-Frequency Radiation ...	139
Rathgeber, H. D.—See Law, P. G., and McKenzie, C. D.		Thermal Radiation from the Moon	63
Residual Effects of Treatments, Estimation of ...	149	Thomas, W. C.—See Brown, R. D., <i>et al.</i>	
Rheological Systems, Second Viscosity Coefficient in ...	394	Thomas, W. C.—See Lahey, F. N.	
Ritchie, E.—See Anet, E. F. L. J., and Hughes, G. K.		Thunderstorm Flight Record ...	550
Ritchie, E.—See Anet, F. A. L., and Hughes, G. K.		Tore, Uniform Flexure of an Incomplete ...	469
Roberts, J. A.—See Bailey, V. A.		Tore, Uniform Torsion of an Incomplete ...	354
Rutaceae: <i>Acronychia Baueri</i> , Alkaloids of Australian	423, 622, 630	Townsend, A. A.—The Fully Developed Turbulent Wake of a Circular Cylinder ...	451
Rutaceae: <i>Evodia xanthoxyloides</i> , Alkaloids of Australian	429	Transients in an Ionized Medium	322
Rutaceae: <i>Melicope fareana</i> , Alkaloids of Australian	249–82	Trikojus, V. M.—See Stanley, P.	
Salicylaldehyde, Metal Complexes of ...	92	Tschu, K. K.—On the Practical Determinations of Lunar and Luni-Solar Daily Variations in Certain Geophysical Data ...	1
Smith, Betty.—See Kraus, E. B.		Turbulence, Diffusion in a Field of Homogeneous ...	437
Smith, E. J.—Experiments in Seeding Cumuliform Cloud Layers with Dry Ice ...	78	Viscosity Coefficient in Rheological Systems ...	394
Smith, E. J.—See Squires, P.		Wave Equations for Electromagnetic Radiation in an Ionized Medium in a Magnetic Field	169
Smith, R. C. T.—See Freiberger, W.		Wavelengths, Bursts of Solar Radiation at Metre ...	214
Solar Radiation at 1200, 600, and 200 Mc/s. ...	48	Wavelengths, 1200, 600, and 200 Mc/s., Solar Radiation at ...	48
Solar Radiation at Metre Wavelengths, Noise-like Character of Solar Radio-Frequency Radiation of Thermal Origin ...	228 198	Waves, Electro-Magneto-Ionic ...	307

	PAGE		PAGE
Westfold, K. C.—		Yabsley, D. E.— <i>See</i> Christiansen,	
The Wave Equations for Elec-		W. N., and Mills, B. Y.	
tromagnetic Radiation in an		Yabsley, D. E.— <i>See</i> Lehany,	
Ionized Medium in a Magnetic		F. J.	
Field ... ..	169	Yabsley, D. E.— <i>See</i> Pawsey,	
Westfold, K. C.— <i>See</i> Jaeger, J. C.		J. L.	
Williams, E. J.—			
Experimental Designs Balanced			
for the Estimation of Residual			
Effects of Treatments ...	149		





DATE DUE			
[REDACTED]			
AUG 20 1983			
RETD AUG 19 2003			
UIC JUN 01 2007			
UIC Rec'd MAY 21 2007			
UIC JUL 6 2009			
UIC Rec'd JUN 25 2009			
PERIODICALS MUST BE RETURNED TO PERIODICALS DESK ONLY			
DEMCO 38-297			

DEMCO 38-297





3 8198 304 260 613

UNIVERSITY OF ILLINOIS AT CHICAGO

4

